Testing for persistence in US mutual funds' performance: a Bayesian dynamic panel model.

Emmanuel Mamatzakis[a] and Mike G. Tsionas[b]

July 2020

We provide a Bayesian panel model to consider persistence in US funds' performance while we tackle the important problem of errors in variables. Our modelling departs from prior strong assumptions such as error terms across funds being independent. In fact, we provide a novel, general Bayesian model for (dynamic) panel data that is stable across different priors as reported from the mapping of the prior to the posterior of the Bayesian baseline model with the adoption of different priors. We demonstrate that our model detects previously undocumented striking variability in terms of performance and persistence across funds categories and over time, and in particular through the financial crisis. The reported stochastic volatility exhibits a rising trend as early as 2003-2004 and could act as an early warning of future crisis.

Keywords: US mutual fund performance, Bayesian panel model time-varying stochastic heteroskedasticity, time-varying covariance.

[a]Birkbeck College, University of London, Department of Management, London WC1E 7HX, UK. E-mail: e.mamatzakis@bbk.ac.uk. [b]Lancaster University Management School, Bailrigg, Lancaster, LA1 4YX, UK, E-mail: m.tsionas@lancaster.ac.uk.

1. INTRODUCTION

Assessment over fund managers' performance has received considerable attention since the seminal paper of Jensen (1968) with mixed findings, as it is openly challenged whether funds would outperform their passive benchmark (Gruber 1996, Carhart 1997; Lunde, et al. 1999; Fama and French 2010; Basak and Makarov, 2014; Cullen et al., 2012; Cabello et al. 2014; Utz et al. 2015; Vidal-García et al. 2018; Giuzio Kay et al. 2018).

At the core of the dispute is accurately measuring the performance of funds. Traditional performance measures compare the returns of the examined portfolio to the returns of an unmanaged portfolio of comparable risk. A number of measures of funds' performance such as the net return ratio, the abnormal return using panel data set (Khorana and Servaes, 2012; Blake, et al. 2014, 2017). The abnormal return is the difference between fund's return and the return of a portfolio which share the same risk characteristics as the fund in consideration. Other measures include: a dummy that equals to 1 if a particular family of funds has at least one fund operating in the top 5% best performing funds of a given category in a given year (Khorana and Servaes, 2012); the Sharpe ratio (Daraio and Simar, 2006), whereas risk is computed as the standard deviation of monthly returns (Huang et al., 2007).[1] Ferson and Lin (2014) using panel data focus on alphas and argue that there should be some bounds that depend on cross sectional investor heterogeneity with the flow response to past fund alphas. This strand of research picks earlier findings (see Clode, 2011; and Busse 2001) arguing that alphas might not be without issues when it comes to select a fund. The underlying

[1] However, several drawbacks of these metrics such as their inability to incorporate funds' transaction costs or the issue of selecting the proper benchmark have fuelled the introduction of performance measures that rely on frontier analysis in the spirit of Koopmans (1951) and Farrell (1957).

autocorrelation could explain results whist the hypothesis of funds being cross-sectionally independent might not be valid (Goriaev et al. 2005).[2]

Beyond issues related with accurately measuring finds' performance, there is an open discussion regarding what are the important covariates of funds' performance. As expected, focus has been for some time on the role of risk. Most studies show that, indeed, risk is important for funds' performance (Giuzio Kay et al. 2018; Vidal-García et al. 2018; Utz et al. 2015; Basak and Makarov, 2014; Brown et al., 1996; Cullen et al., 2012; Goriaev et al., 2005; Koski and Pontiff, 1999).[3] Brown et al. (2001) examine both competition and risk in the hedge fund, reporting similar results as in Brown et al., (1996). Busse (2001) show that poorly performing fund managers alter their risk to be able to catch up with interim winners at the end of the year. Basak and Makarov (2014) focus on the manager's portfolio choice with respect to the strategic interactions among managers competing for fund flows. Their model builds on the strategic behaviours of two risk-averse managers, revealing that a manager either wins or loses, and never opts for a draw.[4]

Other studies (Prather et al., 2004; Vidal-García et al. 2018; Giuzio Kay et al. 2018) report the link between fund performance and various operational characteristics such

[2] Kempf and Ruenzi (2008) study the competition between fund managers across funds' family. They argue that an optimal policy of fund managers is to alter their risk-taking. Studying US equity mutual funds between 1993 and 2001, they report the presence of the family tournament, which is more pronounced in large families.

[3] Brown et al. (1996) identify that interim losers who underperform the benchmark in the first half of the year are likely to increase their risk relative to mid-year winners. Funds are ranked according to their cumulative return, while risk is measured by the ratio of fund's standard deviation after the interim performance assessment to its standard deviation before that date. Another proxy for risk is the tracking error variance, which is the variance of the difference between fund's return and the value-weighted market index (Chevalier and Ellison, 1997).

[4] Basak and Makarov (2014) show that, even when a manager is significantly ahead in the tournament, her investment behaviour and thus portfolio volatility is still influenced by the tournament incentives. In addition, Sato (2015) show the importance of flow-performance relationship and asset bubbles.

as expenses, size, past performance. This type of information could be rather beneficial to investors who decide among offered funds should a reliable relation exists between a fund's performance and some of its observable characteristics. Ferson and Mo (2016) argue that a well-specified performance measure should be based on the sum of covariance between the portfolio holdings and the subsequent abnormal, or risk-adjusted returns, with an underlying stochastic discount factor (see also Cabello et al. 2014). Their modelling has certain appeal, but it still does not address issues related to time-varying covariance where the evidence shows that indeed this is the case (Cabello et al. 2014; Utz et al. 2015; Ferson and Mo 2016; Basak and Makarov 2014; Blake, et al. 2014, 2017).

From the above literature becomes apparent that to date there is no silver bullet regarding an appropriate modelling of mutual fund performance and its underlying determinants across funds and over time. This paper bridges a gap in the literature by providing a novel way modelling mutual funds' performance, relaxing some of the strong assumptions in the literature. Moreover, we argue that time-varying heteroskedasticity and time-varying covariances (in line with Ferson and Mo 2016; Blake, et al. 2014, 2017) are of importance for measuring mutual fund performance without resorting to strong assumptions regarding the unobservable underlying idiosyncratic characteristics of the fund managers. To this end, the purpose of our study is fourfold. First, we propose a new Bayesian panel model that captures time-varying heteroskedasticity and time-varying covariances in funds' performance as well as general autocorrelation and the underlying stochastic volatility. Second, this model allows measuring persistence and it takes also into account errors in the variables. This is commonly acknowledged (Annaert et al. 2003; Barber 2012; Ferson and Mo 2016;

Basak and Makarov 2014; Blake, et al. 2014, 2017), but we are not aware of any previous studies that have dealt with the issue. Our modelling departs from prior strong assumptions, for example that error terms across funds are independent (Ferson and Mo 2016; Basak and Makarov 2014; Blake, et al. 2014, 2017; Casarin and Marin, 2009). Third, as the estimation of this new model is cumbersome, we apply Bayesian techniques that facilitate the robustness of the estimation (Annaert et al. 2003; Barber 2012). Bayesian analysis is implemented using state-of-the-art Sequential Monte Carlo / Particle-Filtering (SMC/PF) techniques. Fourth, we broaden the findings of the relatively few studies measuring fund performance for an up to date set concerning US mutual funds for which we demonstrate that results remain stable across different priors as reported from the mapping of the prior to the posterior.

A preliminary review of our results reveals that risk asserts a positive and significant impact on US mutual finds' performance across different specifications, whilst all Fama and French five factors also show strong positive and significant effect on funds' performance. There has been striking variability in terms of performance and persistence across funds categories and over time, and in particular through the financial crisis. The reported stochastic volatility exhibits a rising trend as early as 2003-2004 and could act as an early warning of future crisis. We show that our results are stable across different priors as reported from the mapping of the prior to the posterior of the Bayesian baseline model with the adoption of different priors. Using likelihood-based techniques, especially Bayesian methods organized around Sequential Monte Carlo, we avoid the need for asymptotically-based inferences which can be misleading in finite samples and in general models which consider almost all features of commonly employed panel data sets.

The rest of the paper is organized as follows: Section 2 presents the new performance model for funds, whilst section 3 reports the data set. Section 4 discusses empirical results. Finally, Section 5 provides some concluding remarks and policy implications.

2. METHODOLOGY

2. 1 Fund's performance and persistence model

So far in the literature performance is characterised by a number of measures, such as net return ratio, abnormal return, Sharpe ratio (Daraio and Simar, 2006; Khorana and Servaes, 2012; Blake, et al. 2014, 2017; Ferson and Mo 2016; Ferson and Lin 2014; Casarin and Marin, 2009). Moreover, we build on the earlier research by Blake, et al. (2014, 2017) suggest that an efficient way of measuring mutual funds' performance is to apply bootstrap methods. The authors effectively pool observations over time, whereas some cross-correlation of fund returns is allowed. We argue that such modelling is too restrictive in the underlying assumptions as heteroskedasticity, errors in variables, covariance across funds, and volatility are not effectively captured. In addition, we augment Ferson and Mo (2016) by allowing time varying heteroskedasticity and covariance of fund performance.

At this point it is, perhaps, of interest to explain why we use a Bayesian approach: (a) we have a number of latent variables in our model so, from the practical point of view, the Bayesian approach is preferable; (b) in the Bayesian approach it is possible to test the effect of various prior assumptions on the results; (c) Bayesian inference provides exact (as opposed to asymptotically—based) results for the given data. In the case of mutual funds, we find this last point as particularly important. The frequentist framework (particularly for *t*-statistics of selectivity) rely on fund performance that could have been observed but it was not actually observed. It is possible to obtain Bayesian results conditional on the fact that certain parameters are (or are not) statistically "significant" (which, again, depends on the universe of data that could have been observed but they never actually did). In dynamic panel data (DPD) the situation is even more critical as existing method of moment techniques (Chen et al., 2004; Ferreira et al., 2012; Chen et al., 2013; Khorana and Servaes, 2012) may behave erratically in finite samples and depend on the use of instruments whose validity has to be tested, although there is no satisfactory testing procedure for valid instruments to the best of our knowledge. It is worth noting that the relevance and strengthens of the frequentist approach is not challenged by our proposed Bayesian approach. Our choice of Bayesian approach is related to its computational feasibility. The only successful Bayesian study on which we rely, for the most part, is Hsiao, et al. (1999) which has not diffused much into the literature on DPD and mutual funds, in particular.

Given the above, we propose the following model:

$$y_{it} = \gamma_{it} + x_{it}'\beta + v_{it},\ i = 1,\dots,n,\ t = 1,\dots,T, \tag{1}$$

where the dependent variable, y_{it}, is the return of mutual fund *i* at date *t*, x_{it}' is a *kx1* vector of covariates (such variables are fund-specific like loads, fee, turnover, expenses, risk and turnover ratio, but also we include Fama-French 5 factors), γ_{it} captures fund- and time-specific effects, that in the context of the present analysis captures '*generalized Jensen's alphas*', while v_{it} is an error term. We intend to propose a *general* model for panel data to allow for structure in returns.

The first novelty that we introduce is that error terms across funds cannot be independent: $v_t = [v_{1t},...v_{nT}] \overset{iid}{\sim} N_n(0,\Sigma),\ t = 1,...,T$. Later on we intend to modify the iid assumption. Secondly, we intend to model the g_{it} in a non-parametric way. This is of importance and complements previous research by Blake, et al. (2014, 2017) that apply bootstrap methods to measure g_{it} with some strong underlying assumptions.

Herein we build on Koop and Poirier (2004), arguing that $\gamma_{it} - \gamma_{i,t-1} \sim N(0,\omega^2)$ which is equivalent to a spline model[5]. It is important to emphasize that this is a Bayesian interpretation of standard non-parametric procedures because it imposes a prior notion about smoothness in the sense that, as a function of time, the γ_{it} s are likely to behave smoothly.

We extend such model as follows:

$$\gamma_{it} - \gamma_{i,t-1} \sim N(\alpha_i, \omega_i^2), \tag{2}$$

where α_i is a '*persistence effect*' across funds and we allow w_i to be also fund-specific. This effect denotes persistence over time in the mutual fund "skill" as it is commonly referred to.

Moreover, we extend below to a model where first derivatives are likely to be smooth as well. Parameters γ_{it} are of major importance here, as we are interested in the performance of mutual funds. Moreover, α_i measure average difference of performance for a mutual fund when all covariates χ_{it} have been set to given values (their means, say).

[5] Initial conditions are treated as unknown parameters.

This perspective allows modelling volatility across funds whilst we also account for persistence in funds' performance at the fund level. Moreover, if we take first differences, we have: $\Delta y_{it} = \alpha_i + \Delta x_{it}'\beta + \Delta v_{it} + \xi_{it},\ i = 1,...,n,\ t = 2,...,T,$ (3),

where $\xi_{it} \sim N\left(0, \omega_i^2\right)$ and Δv_{it} follows an MA(1) process with a unit coefficient (if Σ is diagonal). This, in fact, shows that persistence can be measured using the model in first differences while performance can be measured using directly the g_{it} s.

Finally, we address another important problem that has not received enough attention in the mutual fund evaluation literature. This is the problem of errors in the variables. In addition, we also allow the β's to be time varying. Therefore, we modify the model as[6]:

$$y_{it} = \gamma_{it} + x_{it}^{*\prime}\beta_t + v_{it},\ i = 1,...,n,\ t = 1,...,T, \tag{4}$$

where $x_{it}^{*\prime}$ denotes the actual data, and $x_{it} = x_{it}^{*\prime} + \varepsilon_{it}$

where ε_{it} denotes measurement error.[7]

Given the formulation in (2) it is clear that even if we estimate (4) in first-differences (which we do not) then the time-invariant or persistent effects α_i still appear in the model.

We assume: $$\varepsilon_t = \left[\varepsilon_{1t},...,\varepsilon_{nt}\right]' \sim N_n\left(0, \Omega\right). \tag{5}$$

[6] It is easy to show that the interpretation of γ and α goes through even when βs are time-varying.

[7] Given the model of equation (4) the following applies: i) we can allow for different temporal coefficients, b_t, and ii) we can allow for arbitrary patterns of autocorrelation, since we can assume: $E\left(v_s v_t'\right) = \Sigma_{st},\ s,t \in \{1,...,T\}$. In addition, we allow for arbitrary autocorrelation and arbitrary forms of heteroskedasticity as well. This comes at the cost of allowing for $\frac{T(T+1)}{2}$ matrices of the form Σ_{st} each of which is $n \times n$.

To determine a prior for Ω we use the decomposition $\Omega = C'C$ where C is an upper triangular matrix. Let $c = vec(C)$, where *vec* vectorizes the elements in the upper diagonal. Our prior is:

$$c \sim N(\bar{c}, \bar{V}_c). \qquad (6)$$

For simplifying the notation for a given time period the model in equation (4) can be written as:

$$y_t = X_t\beta_t + \gamma_t + v_t, t = 1,...,T, \qquad (7)$$

where $y_t = \left[y_{1t},...,y_{nt}\right]'$, $X_t = \left[x_{1t}',...,x_{nt}'\right]'$, $\gamma_t = \left[\gamma_{1t},...,\gamma_{nt}\right]'$, and $v_t = \left[v_{1t},...,v_{nt}\right]'$.

If we modify X_t to include an identity matrix with dimension $n \times n$ and expand b_t with a vector of g_{it} with dimension $n \times 1$, the model can be written without loss of generality as: $y_t = X_t\beta_t + v_t, t = 1,...,T.$ (8)

If we define $X = diag\left[X_1,...,X_T\right]$ we can write the model in compact form as:

$$y = \Gamma + X\beta + v, \qquad (9)$$

where $v = \left[v_1',...,v_T'\right]'$, Γ denotes the stacked vector of γ_{it} conformably with y_{it} and

$$E(vv') = \underset{(nT \times nT)}{\Sigma} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \; ... \; \Sigma_{1T} \\ \Sigma_{12} & \Sigma_{22} \; ... \; \Sigma_{2T} \\ & \cdots \\ \Sigma_{1T} & \Sigma_{2T} \; ... \; \Sigma_{nT} \end{bmatrix}. \qquad (10)$$

This is the general form of our model in (7), (8), (9) or (4) in more "accessible" form. The form is useful in that it allows us to define more easily the covariance matrices between errors of different mutual funds but in other respects one can still stick with the simpler form (2) where vector Γ does not appear in (vector) form.

As there are $\frac{nT\,(nT+1)}{2}$ free elements in Σ, there is a huge number of parameters to estimate. We intend to place priors on Σ which tend to favour lack of autocorrelation but allow for heteroskedasticity. This means that Σ_{st} should be *'close'* to a zero $n \times n$ matrix ($s \neq t \in \{1,...,T\}$), say $O_{(n\times n)}$.[8] If we leave Σ_{tt} ($t \in \{1,...,T\}$) unrestricted, this means that we have $\frac{n(n+1)}{2}$ parameters for each ones of the T matrices, which is still excessive.[9]

From our previous discussion on g_{it}s *it is clear that a spline or smoothness prior will essentially result in a non-parametric model*. Thus, for the elements of β_t, the smoothness prior is of the form:

---

[8] Regarding the elements of $\Sigma_{ts}, t \neq s$, since the prior belief is that these are all zero matrices, we can adopt a '*model selection prior*' (see Koop 2013) of the form: $\frac{\sigma_{ij,t}}{\sqrt{\sigma_{ii,t}\sigma_{jj,t}}} = \begin{cases} 0, & \text{with probability } p, \\ \rho, & \text{with probability } 1-p. \end{cases}$

In this prior we set $p = \frac{1}{2}$ and we treat ρ as unknown parameter with a flat prior.

[9] Regarding Σ_{tt} the diagonal elements, say $\sigma_{ii,tt}$, allow for arbitrary time-varying heteroskedasticity while $\sigma_{ij,tt}$ allows for contemporaneous correlation of returns. The matrix has received a great deal of attention in DCC and similar models. Suppose that: $\Sigma_{tt} = H_t H_t'$, where H_t is an $n \times n$ upper diagonal elements and its non-zero elements can be vectorized as: $h_t = vec\left[h_{ij,t}, i, j = 1,..., \frac{n(n+1)}{2}\right]$. We proceed with a prior assuming that: $h_t = a + Ah_{t-1} + u_t$, where α and A have dimension $\frac{n(n+1)}{2} \times 1$ and $\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}$, respectively, and the error term $u_t \sim N(0, \Psi)$. To determine a prior for Ψ we use the decomposition $\Psi = C_\Psi' C_\Psi$ where C_Ψ is an upper triangular matrix. Let $c_\Psi = vec(C_\Psi)$. Our prior is: $c_\Psi \sim N(\bar{c}_\Psi, \bar{V}_\Psi)$. In effect, we place a multivariate stochastic volatility prior. Although we have high dimensional objects α and A, we can proceed with priors to resolve the curse of dimensionality: $a \sim N(\bar{a}, \bar{V}_a), vec(A) \sim N\left(\bar{A}, \bar{V}_A\right)$.

$$\beta_t - \beta_{t-1} \sim N\left(\delta, \Xi\right),\ \Xi = diag[\xi_1^2, ..., \xi_n^2], \tag{11}$$

$$\text{or } \beta_t - 2\beta_{t-1} + \beta_{t-2} \sim N\left(\delta, \Xi\right),\ \Xi = diag[\xi_1^2, ..., \xi_n^2]. \tag{12}$$

We call these models spline-I and spline-II, respectively and our underlying priors are:

$$\delta \sim N(\bar{\delta}, \bar{V}_\delta), \tag{13}$$

$$\log \xi_i^2 \sim N\left(\bar{\varphi}, \bar{V}_\xi\right),\ i = 1, ..., n. \tag{14}$$

Before proceeding we need to mention that the issue of measurement errors has been studied and addressed within a GMM framework (see, for example, Biørn, 2015). Therefore, we do not wish to claim superiority of the Bayesian approach here. Second, the issue of mis-measurement of regressors can be certainly addressed properly within a frequentist framework (Hayakawa and Qi, 2019) so, the issue of the “false" significance can be solved. To conclude, we do not view the Bayesian approach as inherently superior given the relevance and strengthens of the frequentist approach. The Bayesian approach is used here mainly because it is computationally convenient and also because it allows one to examine sensitivity to prior assumptions. This issue does not arise in the frequentist approach but then there are well known issues with use of *p*-values which are documented in the literature (see for example American Statistical Association, statement on *p*-values, 2016) but are beyond the scope of the paper. Computational convenience arises because of the smoothness model in (11) and / or (12). Such models are estimated typically using a Bayesian approach (see, for example the software implementation in BayesX).

2.2 Specification of the baseline prior and its variations

In (7) we have $c \sim N(\bar{c}, \bar{V}_c)$ and we set $\bar{c} = 0, V_c = h_c I$ for $h_c = 1$. Similarly we have $c_\Psi \sim N(\bar{c}_\Psi, \bar{V}_\Psi)$ (see footnote 6) and we set $\bar{c}_\Psi = 0, \bar{V}_\Psi = h_\Psi I$ for $h_\Psi = 1$ and set $\bar{a} = 0, \bar{V}_a = h_a I$ with $h_a = 1$ and $\bar{V}_A = h_A I$ with $h_A = 1$. Thus, for $\bar{A}$ we adopt a

Minnesota-like prior where the elements corresponding to the diagonal are $d_{\bar{A}} \sim (0,1)$ and the others are zero. We start by setting $p = \frac{1}{2}$ (see footnote 6) and we treat r as unknown parameter. The prior of r is taken to be $\rho \sim N_{(-1,1)}\left(0, h_\rho\right)$, $h_\rho = 1$. In (14) we have $\delta \sim N(\bar{\delta}, \bar{V}_\delta)$, where $\bar{d} = 0$ and $\bar{V}_\delta = h_\delta I$ for $h_\delta = 1$.

Finally, in (15) we have $\log \xi_i^2 \sim N\left(\bar{\varphi}, \bar{V}_\xi\right)$, $i = 1, \ldots, n$. We set $\bar{j} = 0$ and $\bar{V}_\xi = h_\xi I$ where $h_\xi = 1$. In the baseline prior, there are many parameters that we can vary to perform sensitivity analysis. To facilitate the analysis, we present the variations of baseline priors in Table 1.

INSERT TABLE 1 HERE

In the empirical application, our intention is to run the baseline model and then adopt different priors as in Table 1 to examine how the results change so that we have some sense of the mapping from the prior to the posterior. To explore the posterior, we use SMC/PF techniques (see Technical Appendix I).

3. DATA

3.1. Sample description

We obtain mutual fund data from Morningstar database for the 2000-2014 period. There are 10,391 funds (94,670 observations), 459 families (5,689 observations), and 25 Morningstar categories (366 observations). Table 2 provides the descriptive statistics at the fund family level for every five years in the sample. Our sample includes US funds in different categories available during the observed period. There are 1623 funds, which charge redemption fee, 1875 funds charging front load, and 260 funds charging both types of fees. Some studies of mutual funds exclude load funds to avoid the problem of addressing different sales fees in the fund's operating costs (Babalos et al., 2015; Gil-Bazo and Ruiz-Verdú, 2009). Our study, in contrast, accounts for both load and no-load funds, and use the information on front-end and back-end loads as fund's characteristic variables as in Daraio and Simar (2006). As indicated in Ferris and Chance (1987) funds' expenses do not comprise load charges. While front-end loads are sales charges paid to brokers or financial advisors for selling the fund, back-end loads are levied on customers for redeeming their shares (Daraio and Simar, 2006; Khorana and Servaes, 2012). The fact that redemption fee exists could inhibit families' competitiveness as investors could be hindered from leaving the funds, especially when funds appear to be underperforming (Khorana and Servaes, 2012). As a result, one may presume that no-load funds attract more investors. However, for those who would need professional advice for their investment choices, front-end load could be a reasonable premium they are willing to compensate for financial advisors (Ferris and Chance, 1987). Additionally, there could be the probability of no-load funds imposing higher other fees on their investors (Tran-Dieu, 2015) or incurring higher expense ratio compared to their load peers (Ferris and Chance, 1987). Hence, analysing this comprehensive sample would produce inclusive results on the competitiveness of different types of funds in the US mutual fund industry.

INSERT TABLE 2 HERE

Based on Morningstar classification, there are 25 fund categories in our sample after being reviewed for errors and outliers. More specifically, according to size, these types consist of large blend, mid-cap blend, small blend, foreign large blend, large growth, mid-cap growth, small growth, foreign large growth, foreign small/mid growth, large value, mid-cap value, and small value. In terms of sector, these categories include real estate, global real estate, technology, equity energy, financial, consumer cyclical, health, utilities, natural resources, communications, consumer defensive, industrials, and world stock.

Regarding variable selection for mutual fund, there is not a universal accepted approach, apart from the return variable that is rather basic. To account for the impact of various determinants we include in addition variables such as loads, fee and turnover. In particular, 12b-1 fee refers to marketing and distribution fee related to money paid to selling agents (Collins and Mack, 1997) and marketing expenses (Khorana and Servaes, 2012). In our sample, there are 7958 funds having a 12b-1 plan. As 12b-1 fee is supposed to be a driving factor in raising fund's assets, there could be two possibilities. On the one hand, economies of scale may exist, which provides funds with the benefit of passing on the fee to both existing and new investors (Khorana and Servaes, 2012). On the other hand, 12b-1 fee can raise expenses as it is a component of a fund's expense (Ferris and Chance, 1987; Latzko, 1999). Front-end load and back-end load have also attracted research interests in their impact on funds' expenses (Daraio and Simar, 2006; Khorana and Servaes, 2012; Latzko, 1999). Front-end loads are an initially one-off sales charge as a reduction to the investment to the fund and are used to incur the cost of financial advisors in attracting new investors. Back-end loads, often known as deferred loads or redemption fee are levied when investors redeem their shares. When redemption fee is high, it may also hinder fund shareholders from leaving the fund, especially underperforming ones.

In subsequent analyses, we also include variables such as total expenses (including loads), risk (measured as the weighted average standard deviation of monthly return), turnover ratio, and number of funds. The literature has suggested that family's diversification across investment styles would benefit investors in terms of fewer restrictions imposed on their asset allocation (Mamaysky and Spiegel, 2002). This would also denote the presence of risk hedging improvement in contrast to the economies of scale arising from '*learning-by-doing externality'*, which exists in more focused families (Massa, 2000). Furthermore, risk is more likely to affect fund family's market share through its relationship with competition and performance (Basak and Makarov, 2012; Huang et al., 2011; Huang et al., 2007; Spiegel and Zhang, 2013; Vidal-García and Vidal, 2014).

The number of funds offered by the fund family is also included to observe whether there is a presence of cost sharing between funds. Put differently, a fund family may enjoy greater economies of scale as the expenses could be reduced for a group of funds (Malhotra et al., 2007). As indicated in Khorana and Servaes (2012), the number of fund started could signify additional business lines, product differentiation, or simply the incentive to increase the likelihood of having funds on the top 5% best-performing classification. It squared value gives an indication for the outstanding impact (if any) in case there is a considerable number of new funds.

Based on Morningstar's definition, turnover ratio conveys the fund's trading activity. Funds report this figure by taking the lesser of purchases or sales of all securities with maturities from one year and dividing by average monthly net assets. The lower the turnover ratio, the more the fund is in favour of the buy-and-hold strategy. Stated differently, high turnover ratio indicates active portfolio management strategies (Daraio and Simar, 2006; Khorana and Servaes, 2012). As a result, active fund managers classified based on high turnover ratios may impose more transaction costs on fund shareholders, in turn raising the fund total expenses.

We also include the Fama-French 5 factors. Moreover the factors are: factor 1 the SMB (Small Minus Big) defined as the average return on the nine small stock portfolios minus the average return on the nine big stock portfolios, factor 2 the HML (High Minus Low) that is the average return on the two value portfolios minus the average return on the two growth portfolios, factor 3 the RMW (Robust Minus Weak) the average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios, factor 4 the CMA (Conservative Minus Aggressive) the average return on the two conservative investment portfolios minus the average return on the two aggressive investment portfolios, and the last factor R_m-R_f, the excess return on the market (see Fama and French, 2014).[10]

4. EMPIRICAL RESULTS

In this section, we run the baseline model and then adopt different priors as in Table 1 to examine how the results change so that we have some sense of the mapping from the prior to the posterior. We do not report results for autocorrelation or correlation across funds and statistics related to errors-in-variables to save space but these are available on request.

For prior sensitivity analysis, we simulate 10,000 priors from the last column of Table 1 and we repeat posterior analysis using SMC/PF techniques (Appendix I).

4.1 Selecting the model

[10] Moreover, the 5 factors come from Fama and French (2014) and defined as SMB(B/M) = 1/3 (Small Value + Small Neutral + Small Growth) - 1/3 (Big Value + Big Neutral + Big Growth), SMB(OP) = 1/3 (Small Robust + Small Neutral + Small Weak) - 1/3 (Big Robust + Big Neutral + Big Weak), SMB(INV) = 1/3 (Small Conservative + Small Neutral + Small Aggressive) - 1/3 (Big Conservative + Big Neutral + Big Aggressive) and thus SMB = 1/3 (SMB(B/M) + SMB(OP) + SMB(INV)). HML =1/2 (Small Value + Big Value) - 1/2 (Small Growth + Big Growth). RMW = 1/2 (Small Robust + Big Robust) - 1/2 (Small Weak + Big Weak). CMA =1/2 (Small Conservative + Big Conservative) - 1/2 (Small Aggressive + Big Aggressive). Data include all NYSE, AMEX, and NASDAQ firms.

To get an idea of what is a '*good*' model we use the baseline prior and we present in Table 3 values of the marginal likelihood (converted into Bayes factors, BF). For a posterior distribution:[11]

$$p(\theta \mid Y) = \frac{L(\theta;Y)\,p(\theta)}{p(Y)}, \qquad (16)$$

the denominator is the marginal likelihood, $p(Y) = \int L(\theta;Y)\,p(\theta)d\theta$ (17).

INSERT TABLE 3 HERE

Clearly, no simplification is possible, at least using the baseline prior as the Bayes factors for alternative models are inferior to the model without restrictions. It also turns out that spline-II behaves much better compared to spline-I.

Our fundamental objective is to evaluate the performance of the mutual funds. From (7) the elements of γ_t are the '*generalized Jensen's alphas*' for all funds for date t. From (2) coefficients α_i denote the persistence. To take off, we would like first of all to present a measure of overall performance, which is

$$PERF_t = n^{-1}\sum_{i=1}^{n}\gamma_{it}, \qquad (18)$$

that is the average performance in the fund industry for a given date.

As the measure depends on all other parameters, in standard Bayesian fashion we take an average across all Monte Carlo draws. Table 4a reports results for model selection, reporting means and standard deviations, including the five factors of Fama and French across specifications. It is perhaps important to point out that our likelihood-based

[11] θ denotes the parameter vector, Y the data and L is the likelihood. Also, p(θ) is the prior.

procedures avoid resorting to asymptotic inferences, which are questionable in finite samples, especially when the econometric model is complicated. Since we use Bayesian techniques we examine thoroughly sensitivity to the prior assumptions as we mentioned before in connection to Table 1.

Table 4a reports that risk asserts a positive and significant impact on performance across specifications, but the model without time-varying β_t. The positive sign comes in line with Basak and Makarov, (2012), Huang et al., (2011), Vidal-García and Vidal, (2014). The convex flow-performance nexus (Chevalier and Ellison, 1997) implies the incentives for fund managers to increase the underlying risk of their funds over time. However, in the model specification where there are no time-varying effects the impact of risk turns negative. Similar results are reported for the third moment, insinuating the importance of allowing for time varying β_t for correctly identifying the impact of underlying moments on performance. Fama and French five factors show strong positive and significant effect on performance. The exception is the RMW factor where negative, but not significant, effects are reported.

In addition, we proceed to a benchmark comparison of model selection (see Table 4a), employing Arellano-Bond-Bover estimator (one step). Indeed, the specification in Table 4a considers the persistence of the dependent variable. Table 4a also reports that expense ratio asserts a significant negative impact on performance across specifications. This is line with Ferreira et al. (2012). Carhart (1997) finds also a negative relationship between fees and net-fee performance. Gil-Bazo and Ruiz-Verdú (2009) opt for pooled ordinary least squares to estimate the effect of funds' expense ratio on before-fee risk-adjusted performance. The results report that before-fee performance is inversely related to fees. According to fund's strategic behaviour, they set fees based on past or expected performance. One underlying rationale is that underperforming funds could have investors who are less sensitive to performance (Christoffersen and Musto, 2002). Therefore, if money does not flow from worse performing funds to the better counterparts, the remained investors would be charged more. Another explanation is that those performance-insensitive investors are target clients of funds with low expected performance (Gil-Bazo and Ruiz-Verdú, 2008). By doing so they are able to charge higher fees, also to compensate for their inability to compete with better performing funds.

INSERT TABLE 4A HERE

Table 4b reports results from the Arellano- Bond-Bover GMM one-step estimator using the internal instruments of first-differenced GMM estimator that consider errors in variables and stochastic volatility. Since the method does not provide SD of returns, first (see first column) we estimate the model without this variable. In a second estimation of the model (see second column), we take variables, including SD of returns, as posterior means from Bayesian estimation and include them as endogenous regressors. The reported results are in line with the ones reported in Table 4a.

INSERT TABLE 4B HERE

Whether size matters for fund performance has attracted much research interest (Chen et al., 2004; Ferreira et al., 2012; Chen et al., 2013; Khorana and Servaes, 2012). Large funds may benefit from economies of scale which they can pass on to investors (Khorana and Servaes, 2012), from investment opportunities which may not be available for small funds, and from better spreads thanks to large positions and trading volumes (Ferreira et al., 2012).[12] This advantage and economies of scope may also be present at the family level (Chen et al., 2004; Chen et al., 2013). Herein we provide statistical significant evidence, under the full model, that size, indeed, matters.[13] The coefficient of fund size takes a positive sign for the full model and the model without cross sectional variation in line with Pollet and Wilson (2008) and Jordan and Riley (2015). However, note that there is variability as results for the models without measurement errors and without time-varying report a negative parameter estimate for fund size, though without significance.

Regarding the turnover ratio, we find that it asserts a positive impact on performance across specification. As turnover ratio indicates the fund family's trading activities, increasing turnover would imply active underlying portfolio management tactics that would also increase performance (Daraio and Simar, 2006; Khorana and Servaes, 2012).[14] Similarly, 12b-1 fees also have a positive and significant effect on performance across specification. 12b-1 fees convey information regarding fund's assets. In this respect, higher fees would imply larger funds that, in turn, we report that have higher performance. In line with economies of scale we show that funds would pass fees to

[12] However, large funds may encounter some disadvantages in terms of liquidity and management (Chen et al., 2004; Pollet and Wilson, 2008). According to the organisational diseconomies hypothesis (Chen et al., 2004), fund size is inversely related to performance. This could be because of hierarchical costs, or management dilution when the fund expands. One manager can easily manage small assets, but it needs a team to manage a large asset base. Large funds would eventually trade large volumes, which may cause difficulties for them in opening and closing their positions. Hence, this liquidity constraint also explains why large funds could be associated with lower performance.

[13] Empirical findings for the size-performance relationship are somewhat mixed. While Chen et al. (2004), Huang et al. (2011) and Ferreira et al. (2012) find a presence of diseconomies of scale, Pollet and Wilson (2008) report that large funds tend to diversify their portfolios which in turn increases their performance. Jordan and Riley (2015) report a positive effect of small size on fund's future alpha. Regressing future alpha on past alpha and size, Elton et al. (2012) do not find a statistically significant predictability of size on future alpha. However, they document that as size increases, expense ratios and management fees decrease.

[14] Funds report turnover ratio by taking the lesser of purchases or sales of all securities with maturities from one year and dividing it by the average monthly net assets. The lower the turnover ratio, the more the fund is in favor of the buy-and-hold strategy. Stated differently, high turnover ratio indicates active portfolio management strategies (Daraio and Simar, 2006; Khorana and Servaes, 2012). As a result, active fund managers classified based on high turnover ratios may induce higher performance.

both existing and new investors as performance rises in line with Khorana and Servaes (2012).[15]

4.2 Funds Performance

The performance results are drawn in Figure 1, based on the full model specification as selected by the results presented in Table 3, along with plus or minus two posterior standard deviations. Clearly, there is a dive in performance during the financial crisis, though the starting point of the downfall is reported as early as in 2006. Thereafter, there is a pick in recovery up till 2012, but that is short lived as a drop is next detected for the remaining of the sample period up till 2014. In this Figure, we detect that funds' performance are following a quite long financial cycle, over ten years period. The slow performance starts as early as in 2001 with the recovery being recorded in 2012. The financial crisis made aggravate things, yet it is evident that funds' financial cycle is elongated. Persistence, therefore, might be of importance here. Next, we report persistence.

INSERT FIGURE 1 HERE

Table 5 reports the average performance indicator across funds categories over the whole sample period. It is striking that some 12 funds report negative performance. The highest performer appears funds in utilities with the lowest performer the real estate funds. Overall performance, as identified also in Figure 1, has been rather subdued during the sample period. There is a dive reported during the financial crisis, but worryingly despite some recovery up till 2012, there is a further decline in recent years.

INSERT TABLE 5 HERE

[15] However, there is some evidence that shows that 12b-1 fee can raise expenses that would eventually compromise performance.

The corresponding marginal posterior density function of persistence according to parameter α_i is shown in Figure 2a. Parameters α_i are fund-specific but time-invariant so we present the posterior distribution of (averaged across Monte Carlo draws) estimates of these parameters. As expected, we have fat tails to the left of the density towards negative values. This is of some interest, as negative performance would persist.[16]

INSERT FIGURE 2A HERE

Funds' posterior mean volatility for 50 different priors, overtime, is presented in Figure 2b. These are filtered estimates from equation (14). In line with the performance results volatility picks in 2009 at the height of the financial crisis and stabilises thereafter. The striking characteristic of our measure of volatility is that is starts picking up as early as 2003-2004, well before the financial crisis. Effectively, our modelling allows measuring risk in early stages and as such could act as early warning.

INSERT FIGURE 2B HERE

In Figure 2c we present marginal posterior densities of $\sqrt{\rho}$ corresponding to cross-sectional correlation for 20 different priors. Evidently, these marginal posteriors are relatively robust and show that $\sqrt{\rho}$ ranges between, roughly, between 0.42 and 0.58 with an average near 0.50.

INSERT FIGURE 2C HERE

[16] Figure of the posterior distribution of α_{is} conditional on a 'significant' γ_{it}, that is the ratio of posterior mean to posterior standard deviation exceeds 2 in absolute value shows that persistence predominantly leans towards negative values (Figure available under request).

To further explore persistence over time, Table 6 reports the average persistence over time. Note that persistence is negative over the whole sample. However during the financial crisis, that is from 2007 to 2009, there was a regime change in funds' performance as reflected by the persistence parameter, which takes positive values. Effectively, we find evidence of a strong negative spiral that further lowers levels of performance during the financial crisis due to strong persistence, though since 2010 persistence is subdued.

INSERT TABLE 6 HERE

Table 7 further reports the average persistence across categories. Once more the dominant fund refers to the utilities with a positive persistence of 0.031 compared to small blend of -0.032.

An interesting question is what happens at the tails of the distribution. A valid but messy way would be to produce posteriors for each fund across all Monte Carlo draws. Instead we provide posterior distributions of posterior mean persistence parameters α_i at different values of the right or left tail (see Figures A1a, A1b and A1c in Appendix II). Fixing, for example, the posterior mean at $a = 0.01$ we take all posterior draws in excess of this value and present the posteriors averaged across all Monte Carlo draws. On the other hand, fixing the posterior mean at $a = -0.01$ we take all posterior draws below this value. Once more, we observe that for negative persistence the densities lean towards negative values while there characterized as leptokurtic compared to positive values for persistence.

4.3 Best and worst fund performers.

Having derived the performance over time and across the main funds categories we turn to the ten best and worst fund performers for an equally weighted portfolio. In a recent Rossello (2015) provide a ranking of investment funds using nonparametric modelling. The results herein complemented previous rankings and extend the modelling techniques.

Moreover, Table 8 reports the ten best and worst performance indicators over the sample period.[17] During the financial crisis even, the best performers turn to negative values, and this lasted till 2011. Since then there is a remarkable recovery, beyond the previous pick of the highest performance in 2005. Similarly, the 10 worst performers were hit by the financial crisis, much more harshly compared to the best 10 performers. However, when it comes to the worst performers, it appears that the recovery in 2014 is much stronger than the best performers though for the former volatility is clearly an issue.

INSERT TABLE 8 HERE

On average, see Table 8, performance is close to zero for best performers but its posterior range is, roughly, between -4% and 8%. For the worst performers, the average is close to -4% but ranges, roughly, from -12% to 5% and is clearly shifted to the left compared to the posterior density of best performers.

[17] For completeness we present in Table A1 in Appendix II the performance of ten average funds. In addition, in Appendix II we provide figures for posterior distributions for 10 best/worst funds as well as 10 average funds. Note that we also include figure A2d that reports posterior distributions of for the best-performing fund for years 2000, 2008, 2012 and 2014

In Figure 3 we present the posterior distribution of g_{it} for an equally weighted portfolio of the ten best and ten worst funds (in terms of simple average returns in the sample).[18] Clearly the worst performers are shifted to the left of the density of the best performers.

INSERT FIGURE 3 HERE

4.4 The fund of funds portfolio

Following from Basak and Makarov (2014) on the manager's portfolio choice where they argue that managers either win or lose we propose to model a fund of funds portfolio as robustness for the above findings. Moreover, we propose to solve a simple quadratic programming problem a la Markowitz:

$$\min_{w \in R^n} : w'\mu - A\,w'\Sigma w, s.t. w \geq 0, w'\mathbf{1} = 1 \qquad (18)$$

where, in familiar notation, μ and Σ are taken from parameter estimates in (9).

We consider a fund-of-funds portfolio P that is formed by using the optimal weights w assuming that $A = 3$. Table 8 presents the performance and persistence of the optimal portfolio P.[19] It is striking that performance in a la Markowitz portfolio (see Table 9) is well below the ten best fund performers, whilst persistence is also subdued compared to above results.[20]

INSERT TABLE 9 HERE

[18] In addition, results are available under request of posterior distributions obtained through 20 different priors from the last column in Table 1 to examine prior sensitivity. These results provide evidence of the performance for the ten best (worst) funds and are in line with the evidence herein.
[19] The posterior distribution of optimal portfolio P is presented in Figure A3a and the posterior distribution of average persistence of its component returns is presented in Figure A3b, for 50 different priors from the last column in Table 1. In Figure A3c, we report the posterior distribution of its persistence only when average persistence of its component returns is '*significant*' (viz. the ratio of posterior mean to posterior S.D. exceeds 2 in absolute value).
[20] In Figure A4 in Appendix II we report the posterior distributions of α_i for each fund only when its average γ_i turns out significant, which is an analogue of Figure A2b but on a fund basis this time.

5. PRIOR SENSITIVITY ANALYSIS

An interesting question is what different priors imply about the funds' ability to perform better than the market. In Figure 4, we consider all 10,000 prior distribution assumptions reported in the last column of Table 1, and we present posterior estimates of temporal averages of g_{it} whose $\tau = \frac{E(\bar{\gamma}_i)}{\sqrt{\text{var}(\bar{\gamma}_i)}}$, $\bar{\gamma}_i = T^{-1}\sum_{t=1}^{T}\gamma_{it}$ exceeds 2 or is lower than -2, for at least ten funds. This presents direct evidence as to whether there are priors, which support the idea that funds perform better than the market.

INSERT FIGURE 4 HERE

Overall, the empirical results show that the baseline model and the adoption of different priors from Table 1 show that results are stable as reported from the mapping of the prior to the posterior.

In this paper we shed new light into the performance of funds while robustness and sensitivity analysis show that results are valid through a plethora of alternative specifications/models and priors. Given the robustness of our findings we perceive that there are policy implications for all participants in fund industry such as shareholders, fund managers and financial regulators. In some detail, results show that fund's size would enhance performance, though there is some evidence of variability across models. Shareholders and investors should note that larger funds would benefit from economies of scale type of effects and would drive to higher returns. This result is in line with Pollet and Wilson (2008). Another significant finding refers to the positive impact of risk on performance across most specifications/model (Basak and Makarov, 2012; Huang et al., 2011; Vidal-García and Vidal, 2014). A novel characteristic of our analysis is that examines whether the underlying relationship between risk and performance is subject to variability as indeed we report it is the case. We show that shocks due to risk is key for understanding the underlying reasons for periods of financial turbulence. Regulators and policy makers alike should take note of this finding. In particular, our measure of volatility was picking up the financial crisis as early as in 2004. This measure of volatility could act as an early warning that could be useful for regulators and supervisory authorities, aiming to intervene in the case of excessive risk that would dwindle financial stability.

6. CONCLUSIONS

This paper proposes a novel panel data model so as to capture time-varying heteroskedasticity, time-varying covariances of performance of US funds. Such modelling also permits general autocorrelation plus errors in underlying variables. Our results show that there has been striking variability across funds categories in terms of performance and persistence. Risk asserts a positive and significant impact on performance across different specifications. All Fama and French five factors show strong positive and significant effect on funds' performance. The exception is the RMW factor that turns negative, but it is not significant. There has been striking variability in terms of performance and persistence across funds categories and over time, and in particular through the financial crisis. Persistence in performance during the financial crisis is strong, insinuating a negative spiral to further lower levels of performance. The reported stochastic volatility exhibits a rising trend as early as 2003-2004 and could act as an early warning of future crisis. Finally, we show that our results are stable across different priors as reported from the mapping of the prior to the posterior of the Bayesian baseline model with the adoption of different priors.

The reported results have some implications for investors, professional managers and regulators. The revealed relationships could be part of investors' information set when select a fund whereas fund managers could benefit from the knowledge of components that enhance their portfolio performance. Finally, regulators and supervisory authorities whose task is to safeguard a secure and well-functioning financial system may consider that risk improves performance.

REFERENCES

American Statistical Association, (2016). Statement on *p*-values, Context, Process, and Purpose, ASA, Volume 70, Issue 2.
Andrieu, C. and G.O. Roberts. (2009) The pseudo-marginal approach for efficient computation. Ann. Statistics, 37, 697–725.
Andrieu, C., A. Doucet, and R. Holenstein (2010). Particle Markov chain Monte Carlo methods, Journal of the Royal Statistical Society: Series B, 72, 269–342.
Annaert, J., J. Van Den Broeck and R. Vander Vennet, (2003). Determinants of mutual fund underperformance: a Bayesian stochastic frontier approach. European Journal of Operational Research 151, 617-632.
Basak, S. and D. Makarov, (2014). Strategic asset allocation in money management. The Journal of Finance 69, 179-217.
Barber (2012). Bayesian Reasoning and Machine Learning, Cambridge University Press.
Blake, D., Caulfield, T., Ioannidis, C. and Tonks, I., (2014). Improved inference in the evaluation of mutual fund performance using panel bootstrap methods. Journal of Econometrics, 183 (2), pp. 202-210.
Blake, D., Caulfield, T., Ioannidis, C. and Tonks, I., (2017). New Evidence on Mutual Fund Performance: A Comparison of Alternative Bootstrap Methods. Journal of Financial and Quantitative Analysis, Volume 52, Issue 3, June, pp. 1279-1299.
Berk, J.B. and R.C. Green, (2004). Mutual Fund Flows and Performance in Rational Markets. Journal of Political Economy 112, 1269-1295.
Biørn, E., (2015). Panel data dynamics with mis-measured variables: modeling and GMM estimation, Empirical Economics, vol. 48, 517-535.
Brown, K.C., W.V. Harlow and L.T. Starks, (1996). Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. Journal of Finance, 85-110.
Brown, S.J. and W.N. Goetzmann, J. Park, (2001). Careers and survival: Competition and risk in the hedge fund and CTA industry. The Journal of Finance 56, 1869-1886.
Busse, J.A., (2001). Another look at mutual fund tournaments. Journal of Financial and Quantitative Analysis 36, 53-73.
Cabello, J. M., Ruiz, F., Pérez-Gladish, B., & Méndez-Rodriguez, P. (2014). Synthetic indicators of mutual funds' environmental responsibility: An application of the reference point method. European Journal of Operational Research, 236(1), 313–325.
Carhart, M.M., (1997). On persistence in mutual fund performance. The Journal of finance 52, 57-82.
Cappé, O., E. Moulines, E., and T. Rydén (2005). Inference in Hidden Markov Models. Springer, Berlin.
Casarin, R. and Marin, J.-M, (2009). Online data processing: Comparison of Bayesian regularized particle letters, Electronic Journal of Statistics, 3, 239-258.
Chen, H. and G.G. Pennacchi, (2009). Does prior performance affect a mutual fund's choice of risk? Theory and further empirical evidence. Journal of Financial and Quantitative Analysis 44, 745-775.
Chen, J., H. Hong, M. Huang, J.D. Kubik, (2004). Does Fund Size Erode Performance? Organizational Diseconomies and Active Money Management. American Economic Review 94, 1276-1302.
Chen, J., H. Hong, W. Jiang and J.D. Kubik, (2013). Outsourcing mutual fund management: firm boundaries, incentives, and performance. The Journal of Finance 68, 523-558.

Chevalier, J. and G. Ellison, (1997). Risk Taking by Mutual Funds as a Response to Incentives. The Journal of Political Economy 105, 1167-1200.
Christoffersen, S.E. and D.K. Musto, (2002). Demand curves and the pricing of money management. Review of Financial Studies 15, 1499-1524.
Chopin, N. (2002). A sequential particle filter method for static models, Biometrika 89 (3) 539-551.
Chopin, N. (2004). Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. Annals of Statistics 32, 2385-2411.
Cremers, K.M. and A. Petajisto, (2009). How active is your fund manager? A new measure that predicts performance. Review of Financial Studies 22, 3329-3365.
Cullen, G., D. Gasbarro, G.S. Monroe and J.K. Zumwalt, (2012). Changes to mutual fund risk: Intentional or mean reverting? Journal of Banking & Finance 36, 112-120.
Elton, E.J., M.J. Gruber and C.R. Blake, (2012). Does mutual fund size matter? The relationship between size and performance. Review of Asset Pricing Studies, ras001.
Daraio, C. and L. Simar, (2006). A robust nonparametric approach to evaluate and explain the performance of mutual funds. European Journal of Operational Research 175, 516-542.
De Jonghe, O. and R. Vander Vennet, (2008). Competition versus efficiency: What drives franchise values in European banking? Journal of Banking & Finance 32, 1820-1835.
Doucet, A., N. de Freitas, and N. Gordon (2001). Sequential Monte Carlo Methods in Practice. New York: Springer.
Durham, G., and J. Geweke, (2013). Adaptive sequential posterior simulators for massively parallel computing environments. Working Paper.
Fama and French, (2014). A Five-Factor Asset Pricing Model. Working Paper. Available at SSRN: http://ssrn.com/abstract=2287202.
Ferreira, M.A., A. Keswani, A.F. Miguel, S.B. Ramos, (2012). The determinants of mutual fund performance: A cross-country study. Review of Finance, rfs013.
Fearnhead, P. (2007). Computational methods for complex stochastic systems: a review of some alternatives to MCMC. Statistics and Computing, 18(2):151-171.
Flury, T., and N. Shephard, (2011) Bayesian inference based only on simulated likelihood: particle filter analysis of dynamic economic models, Econometric Theory 27, 933-956.
Ferson, W. and Haitao Mo. (2016). Performance Measurement with Selectivity, Market and Volatility Timing. Journal of Financial Economics, Volume 121, Issue 1, July, Pages 93-110.
Ferson, W. and Jerchern Lin, (2014). Alpha and Performance Measurement: The Effects of Investor Disagreement and Heterogeneity. Journal of Finance, 69(4), 1565-1596.
Fu, X.M., Y.R. Lin and P. Molyneux, (2014). Bank competition and financial stability in Asia Pacific. Journal of Banking & Finance 38, 64-77.
Gilks, W.R. and C. Berzuini, (2001). Following a moving target: Monte Carlo inference for dynamic Bayesian models, Journal of the Royal Statistical Society B 63, 127-146.
Giuzio Kay, M., Eichhorn-Schott, Sandra Paterlini, and Vincent Weber (2018). Tracking hedge funds returns using sparse clones, Annals of Operations Research, July, Volume 266, Issue 1–2, pp 349–371.
Gaspar, J.M., M. Massa, and P. Matos, (2006). Favoritism in mutual fund families? evidence on strategic cross-fund subsidization. The Journal of Finance 61, 73-104.
Gilks, W.R. and C. Berzuini, (2001). Following a moving target: Monte Carlo inference for dynamic Bayesian models, Journal of the Royal Statistical Society B, 63, 127-146.

Gil-Bazo, J. and P. Ruiz-Verdú, (2008). When cheaper is better: Fee determination in the market for equity mutual funds. Journal of Economic Behavior & Organization 67, 871-885.

Gil-Bazo, J. and P. Ruiz-Verdú, (2009). The relation between price and performance in the mutual fund industry. The Journal of Finance 64, 2153-2183.

Glode, Vincent, (2011). Why mutual funds underperform, Journal of Financial Economics 93, 546-559.

Goriaev, A., T.E. Nijman, B.J. Werker, (2005). Yet another look at mutual fund tournaments. Journal of Empirical Finance 12, 127-137.

Gordon, N.J. (1997). A hybrid bootstrap filter for target tracking in clutter. IEEE Transactions on Aerospace and Electronic Systems 33: 353– 358.

Gordon, N.J., D.J. Salmond, and A.F.M. Smith (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEEE-Proceedings-F 140: 107–113.

Gruber, M.J., (1996). Another puzzle: The growth in actively managed mutual funds. The Journal of Finance 51, 783-810.

Guercio, D.D. and J. Reuter, (2014). Mutual fund performance and the incentive to generate alpha. The Journal of Finance 69, 1673-1704.

Hall, J., M.K. Pitt, and R. Kohn (2014). Bayesian inference for nonlinear structural time series models. Journal of Econometrics 179 (2), 99–111.

Hayakawa, K. and Qi, M., (2019). Further Results on the Weak Instruments Problem of the System GMM Estimator in Dynamic Panel Data Models, Oxford Bulletin of Economics and Statistics, forthcoming.

Huang, J., C. Sialm and H. Zhang, (2011). Risk shifting and mutual fund performance. Review of Financial Studies, Volume 24, Issue 8, August, Pages 2575–2616.

Huang, J., K.D. Wei, H. Yan, (2007). Participation costs and the sensitivity of fund flows to past performance. The Journal of Finance 62, 1273-1311.

Hsiao, Chen, Hashem Pesaran, M., and Kamil Tahmiscioglu, A.. 1999. Bayes Estimation of Short-Run Coefficients in Dynamic Panel Data Models. In Chen Hsiao,

In, F., M. Kim, R.J. Park, S. Kim, T.S. Kim, (2014). Competition of socially responsible and conventional mutual funds and its impact on fund performance. Journal of Banking & Finance 44, 160-176.

Jordan, B.D. and T.B. Riley, (2015). Volatility and mutual fund manager skill. Journal of Financial Economics 118, 289-298.

Kempf, A. and S. Ruenzi, (2008). Tournaments in mutual-fund families. Review of Financial Studies 21, 1013-1036.

Kempf, A., S. Ruenzi and T. Thiele, (2009). Employment risk, compensation incentives, and managerial risk taking: Evidence from the mutual fund industry. Journal of Financial Economics 92, 92-108.

Keswani, A. and D. Stolin, (2006). Mutual Fund Performance Persistence and Competition: A Cross-sector Analysis Journal of Financial Research, 29, 349-366.

Keswani, A. and D. Stolin, (2008). Which money is smart? Mutual fund buys and sells of individual and institutional investors. The Journal of Finance 63, 85-118.

Khorana, A., H. Servaes, (2012). What drives market share in the mutual fund industry? Review of Finance 16, 81-113.

Koop G. and Poirier, D. J., (2004). Bayesian variants of some classical semi-parametric regression techniques, Journal of Econometrics, Volume 123, Issue 2, December, Pages 259-282.

Koop, Gary, (2013). Forecasting with medium and large Bayesian VARs. Journal of Applied Econometrics 28, 177-203.

Koski, J.L., J. Pontiff, (1999). How Are Derivatives Used? Evidence from the Mutual Fund Industry. Journal of Finance 54, 791-816.
Lou, D., (2012). A flow-based explanation for return predictability. Review of Financial Studies 25, 3457-3489.
Liu, J. and M. West (2001). Combined parameter and state estimation in simulation-based filtering. In: Doucet, A., de Freitas, N., Gordon, N. (Eds.), Sequential Monte Carlo Methods in Practice. Springer-Verlag.
Lunde, A., A. Timmermann, and Blake D., (1999). The Hazards of Mutual Fund Underperformance: A Cox Regression Analysis. Journal of Empirical Finance, 6, 121-52.
Malhotra, D., R. Martin, P. Russel, (2007). Determinants of cost efficiencies in the mutual fund industry. Review of Financial Economics 16, 323-334.
Mamaysky, H. and M. Spiegel, (2002). A theory of mutual funds: Optimal fund objectives and industry organization.
Murthi, B., Y.K. Choi, P. Desai, (1997). Efficiency of mutual funds and portfolio performance measurement: A non-parametric approach. European Journal of Operational Research 98, 408-418.
Nemeth, C. and P. Fearnhead (2014). Particle Metropolis adjusted Langevin algorithms for state-space models. Pre-print arXiv:1402.0694v1.
Pesaran H., Kajal Lahiri and Lung Fei Lee (eds), Analysis of Panels and Limited Dependent Variables, chapter 11, 268–296. New York, NY: Cambridge University Press
Pitt, M.K., R.S. Silva, P. Giordani, and R. Kohn (2012). On some properties of Markov chain Monte Carlo simulation methods based on the particle filter, Journal of Econometrics, 171(2), 134–151.
Pollet, J.M. and Wilson, M., (2008). How does size affect mutual fund behavior? The Journal of Finance 63, 2941-2969.
Ristic, B., S. Arulampalam, and Gordon N., (2004). Beyond Kalman Filters: Particle Filters for Applications. Norwood, MA: Artech House.
Rossello, D. (2015). Ranking of investment funds: Acceptability versus robustness, European Journal of Operational Research, Volume 245, Issue 3, 16 September, Pages 828-836.
Sato, Y., (2015). Fund Tournaments and Asset Bubbles. Review of Finance, rfv044.
Sharpe, W.F., (1966). Mutual fund performance. Journal of Business, 119-138.
Spiegel, M. and H. Zhang, (2013). Mutual fund risk and market share-adjusted fund flows. Journal of Financial Economics 108, 506-528.
Taylor, J., (2003). Risk-taking behavior in mutual fund tournaments. Journal of Economic Behavior & Organization 50, 373-383.
Utz S., M. Wimmer, R. E. Steuer, (2015) Tri-criterion modelling for constructing more-sustainable mutual funds, European Journal of Operational Research, Volume 246, Issue 1, 1 October, Pages 331-338.
Vidal-García, Javier, Marta Vida, Sabri Boubaker and Majdi Hassan (2018). The efficiency of mutual funds, Annals of Operations Research, Volume 267, Issue 1–2, pp 555–584.
Wahal, S. and Wang, A.Y. (2011). Competition among mutual funds. Journal of Financial Economics 99, 40-59.

Table 1: Variations of the baseline prior.

Parameters and baseline priors	Reference values
$c \sim N(\bar{c}, \bar{V}_c)$, $\bar{c} = 0,\ V_c = h_c I$ for $h_c = 1$.	$c \sim U(-5,5), h_c \sim U\left(0.1,10\right)$
$a \sim N(\bar{a}, \bar{V}_a), vec(A) \sim N\left(\bar{A}, \bar{V}_A\right).$ $\bar{a} = 0, \bar{V}_a = h_a I$ with $h_a = 1$. $\bar{V}_A = h_A I$ with $h_A = 1$.	$\bar{a} \sim U(-5,5),\ h_a, h_A \sim U\left(0.1,10\right)$
$\bar{A}$ Minnesota-like prior, diagonal elements $d_{\bar{A}} \in (0,1]$ otherwise zero.	$d_{\bar{A}} \sim U(0,1)$
$p = \frac{1}{2},\ \rho \sim N_{(-1,1)}\left(0, h_\rho\right), h_\rho = 1$	$p \sim U(0,1), h_\rho \sim U\left(0.1,10\right)$
$\delta \sim N(\bar{\delta}, \bar{V}_\delta),$ $\bar{d} = 0$ $\bar{V}_\delta = h_\delta I$ for $h_d = 1$.	$\bar{\delta} \sim U(-5,5), h_\delta \sim U\left(0.1,10\right)$
$\log \xi_i^2 \sim N\left(\bar{\varphi}, \bar{V}_\xi\right),\ i = 1,\ldots,n.$ $\bar{f} = 0$ $\bar{V}_\xi = h_\xi I$ where $h_\xi = 1$.	$\bar{\varphi} \sim U(-5,5), h_\xi \sim U\left(0.1,10\right)$

Notes: c =vec(C), where vec vectorizes the elements in the upper diagonal C from the equation of measurement errors (6). $a \sim N(\bar{a}, \bar{V}_a), vec(A) \sim N\left(\bar{A}, \bar{V}_A\right).$ captures the persistence effect as in equation (2). ρ is unknown parameter explained in footnote 8. $\delta \sim N(\bar{\delta}, \bar{V}_\delta),$ notes underlying priors in equation (3) and $\log \xi_i^2 \sim N\left(\bar{\varphi}, \bar{V}_\xi\right),\ i = 1,\ldots,n.$ notes underlying priors in equation (15), while $\bar{A}$ is a Minnesota-like prior. $U(a,b)$ denotes the uniform distribution in the interval $\left(a,b\right)$, $b > a$.

Table 2: Summary statistics for funds.

	2000-2004	2005-2009	2010-2014
Total assets (mil USD)	4860	7220	9670
No of families	459	459	459
No of funds per family	23.876	26.792	21.663
Number of funds started	5.710	4.241	4.550
Number of categories	2.617	2.998	2.501
Gross expense ratio	1.308	1.712	1.089
Gross return	5.373	4.649	11.736
Turnover ratio	0.514	0.749	0.505
Risk	4.336	4.215	3.265
12b-1 fee	0.362	0.362	0.362
Max front load	5.145	5.145	5.145
Redemption fee	1.640	1.640	1.640

Notes: This Table reports the mean of variables over five years intervals. Number of funds started is the number of new funds started in a given year. Number of categories represents the different investment objectives within a family. Turnover is the weighted average turnover across all funds in a family. Risk is measured as the weighted average standard deviation of monthly return. Gross expense ratio, gross return, turnover ratio, risk, 12b-1 fee, max front load, and redemption fee are the weighted average values across all funds in a family.

Table 3: The model comparison: Bayes factors.

	Bayes Factors
Formulation in (12), spline-I	
No restrictions	1.000
$b_t = b_o$ are the same	0.0061
$A=0$, no stochastic volatility	0.0056
$b_t = b_o$ and $A=0$	0.0011
No cross-sectional correlation	0.0003
No measurement error	0.0001
Formulation in (13), spline-II	
No restrictions	7.3221
$b_t = b_o$ are the same	0.0414
$A=0$, no stochastic volatility	0.0032
$b_t = b_o$ and $A = 0$	0.0017
No cross-sectional correlation	0.0002
No measurement error	0.0000

Note: The table presents Bayes factors relative to the full model without any restrictions. The Bayes factors are computed using the marginal likelihood from the Sequential Monte Carlo (particle filtering) approach described in Appendix I.

Table 4a: Model selection, posterior means and posterior standard deviations.

	full model		without measurement error		without time-varying b_t		without cross-sectional correlation	
	Post. mean	Post. SD	Post. mean	Post. SD	Post. mean	Post. SD	Post. mean	Post. SD
SD of returns	0.0171	0.0015	0.0031	0.0021	-0.0022	0.0010	0.0134	0.0019
Skewness	0.0043	0.0035	-0.0030	0.0011	-0.0044	0.0014	0.0035	0.0023
Kurtosis	0.0031	0.00012	0.00132	0.0254	0.00441	0.0165	0.012	0.00024
Expense ratio	-0.0014	0.00032	-0.0133	0.0043	-0.0033	0.0017	-0.0144	0.00017
Loads & Turnover	-0.0015	0.00013	-0.0126	0.0032	-0.013	0.0083	-0.0033	0.00012
12b-1 fees	0.0045	0.00017	0.0023	0.0020	0.0033	0.0432	0.0011	0.00044
Net asset value	0.0032	0.00021	0.0011	0.0017	0.0025	0.0019	0.0015	0.00035
Fund size	0.0015	0.00017	-0.0013	0.0015	-0.0022	0.0017	0.0036	0.00026
Net-exp. ratio	-0.0033	0.00024	0.0045	0.0032	0.00210	0.0013	0.0032	0.00323
Turnover ratio	0.0017	0.00015	0.0045	0.0033	0.0037	0.0022	0.0034	0.00021
F&F 1	0.281	0.032	0.171	0.0117	0.128	0.035	0.155	0.054
F&F 2	0.048	0.032	0.023	0.024	0.032	0.0156	0.0245	0.0272
F&F 3	-0.015	0.011	0.0056	0.0137	0.0044	0.0031	-0.0221	0.0217
F&F 4	0.129	0.0522	0.0171	0.0081	0.0225	0.0152	0.2513	0.0351
F&F 5	0.083	0.0714	0.0331	0.0266	0.0742	0.0553	0.0484	0.0493
Bayes R^2	0.9345	0.0454	0.8716	0.0325	0.8652	0.0352	0.8751	0.0255

Note: We have used the prior that out of 10,000 alternative priors in the last column of Table 3 yields the highest value for the marginal likelihood. So, we use the prior that best '*fits*' the data. The five Fama-French (F&F) factors are also included: F&F 1 the SMB (Small Minus Big), F&F 2 the HML (High Minus Low), F&F 3 the RMW (Robust Minus Weak), F&F 4 the CMA (Conservative Minus Aggressive), and F&F 5 R_m-R_f. We also compute a Bayesian variant of R^2, which is computed, using the correlation of actual and predicted values, averaged over all posterior draws. SD implies standard deviation. Post. implies posterior. We have used 15000 iterations of the Sequential Monte Carlo (particle filtering) approach discarding the first 5000 to mitigate possible start up effects. We used 10000 particles per iteration.

Table 4b: Model selection: Arellano- Bond-Bover GMM one-step estimator.

	Without SD of returns		With SD of returns	
	Post. mean	Post. SD	Post. mean	Post. SD
SD of returns			0.0015	0.0002
Skewness	0.0032	0.0024	0.0021	0.0040
Kurtosis	-0.0014	0.0132	0.0014	0.0133
Expense ratio	0.0012	0.0014	-0.0017	0.0012
Loads & Turnover	0.022	0.017	0.0156	0.0044
12b-1 fees	0.032	0.011	0.0010	0.0002
Net asset value	0.014	0.0003	0.0151	0.003
Fund size	0.0019	0.0011	-0.0032	0.0012
Net-exp. ratio	-0.0012	0.0007	0.0051	0.0010
Turnover ratio	0.0021	0.00013	0.0144	0.0012
F&F 1	0.017	0.0030	0.007	0.007
F&F 2	0.020	0.007	0.005	0.003
F&F 3	-0.015	0.004	0.0149	0.0156
F&F 4	0.013	0.002	0.0221	0.0130
F&F 5	0.023	0.002	0.0087	0.0066
R^2	0.217		0.298	

Note: Arellano- Bond-Bover GMM one-step estimator. The five Fama-French (F&F) factors are: F&F 1 the SMB (Small Minus Big), F&F 2 the HML (High Minus Low), F&F 3 the RMW (Robust Minus Weak), F&F 4 the CMA (Conservative Minus Aggressive), and F&F 5 R_m-R_f. SD implies standard deviation. Post. implies posterior.

Table 5: Average performance indicator across funds' categories.

Category	Performance	Category	Performance
Large Blend	0.007 (0.034)	Technology	-0.003 (0.005)
Mid-Cap Blend	0.001 (0.021)	Financial	-0.004 (0.002)
Small Blend	-0.003 (0.012)	Consumer Cyclical	-0.007 (0.012)
Large Growth	0.005 (0.013)	Equity Energy	0.005 (0.012)
Mid-Cap Growth	0.002 (0.007)	World Stock	-0.003 (0.003)
Small Growth	-0.003 (0.004)	Global Real Estate	-0.075 (0.120)
Foreign Large Growth	0.005 (0.002)	Consumer Defensive	0.003 (0.004)
Foreign Small/Mid Growth	-0.003 (0.001)	Real Estate	-0.082 (0.118)
Foreign Large Blend	0.002 (0.004)	Communications	-0.005 (0.017)
Large Value	0.009 (0.012)	Health	0.003 (0.002)
Mid-Cap Value	0.004 (0.08)	Industrials	0.005 (0.015)
Small Value	-0.001 (0.02)	Natural Resources	-0.003 (0.002)
Utilities	0.0151 (0.01)		

Notes: For each category, the Sequential Monte Carlo / Particle Filtering draws for the performance indicator (PERF) are averaged. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration. Posterior standard deviations appear in parentheses.

Table 6: The yearly average persistence over time.

Year	Mean	S.D.	Min	Max
2000	-0.042	0.017	-0.064	0.0032
2001	-0.051	0.022	-0.071	0.0049
2002	-0.055	0.021	-0.075	0.0083
2003	-0.042	0.025	-0.077	0.015
2004	-0.038	0.020	-0.082	0.017
2005	-0.021	0.019	-0.079	0.025
2006	-0.035	0.017	-0.081	0.027
2007	0.032	0.021	-0.087	0.012
2008	0.043	0.025	-0.091	0.007
2009	0.023	0.027	-0.088	0.003
2010	-0.032	0.031	-0.085	0.019
2011	-0.036	0.034	-0.082	0.022
2012	-0.039	0.035	-0.077	0.025
2013	-0.041	0.031	-0.078	0.021
2014	-0.043	0.027	-0.081	0.019
Average	-0.0251	0.0248	-0.0799	0.0152

Note: The persistence parameters α_i are fund-specific but time-invariant so we present the posterior distribution of (averaged across Monte Carlo draws) estimates of these parameters.

Table 7: The average persistence indicator across funds categories.

Category	Persistence	Category	Persistence
Large Blend	0.017 (0.009)	Technology	0.022 (0.017)
Mid-Cap Blend	0.005 (0.002)	Financial	-0.032 (0.011)
Small Blend	-0.032 (0.014)	Consumer Cyclical	0.006 (0.003)
Large Growth	0.012 (0.008)	Equity Energy	-0.005 (0.008)
Mid-Cap Growth	-0.005 (0.003)	World Stock	-0.004 (0.031)
Small Growth	-0.009 (0.002)	Global Real Estate	-0.017 (0.003)
Foreign Large Growth	0.017 (0.003)	Consumer Defensive	0.007 (0.004)
Foreign Small/Mid Growth	0.005 (0.003)	Real Estate	-0.015 (0.005)
Foreign Large Blend	0.017 (0.008)	Communications	-0.043 (0.032)
Large Value	0.012 (0.007)	Health	0.005 (0.004)
Mid-Cap Value	0.005 (0.004)	Industrials	-0.032 (0.025)
Small Value	0.001 (0.001)	Natural Resources	0.005 (0.004)
Utilities	0.031 (0.017)		

Notes: For each category, the Sequential Monte Carlo / Particle Filtering draws for the persistence indicator (α_i) are averaged. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration. Posterior standard deviations appear in parentheses.

Table 8: The Yearly average 10 best and worst performance indicators over time.

	Best				Worst			
Year	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
2000	0.085	0.077	-0.062	0.251	0.003	0.127	-0.241	0.192
2001	0.072	0.055	-0.031	0.182	0.005	0.125	-0.239	0.233
2002	0.079	0.082	-0.071	0.191	0.007	0.119	-0.255	0.241
2003	0.081	0.093	-0.115	0.257	0.0155	0.118	-0.217	0.255
2004	0.067	0.085	-0.101	0.229	0.0221	0.121	-0.221	0.26
2005	0.083	0.079	-0.082	0.212	0.0203	0.031	-0.052	0.087
2006	0.081	0.072	-0.061	0.225	0.0151	0.029	-0.036	0.019
2007	0.065	0.082	-0.105	0.203	-0.0915	0.103	-0.331	0.085
2008	-0.055	0.095	-0.132	0.103	-0.127	0.155	-0.567	0.061
2009	-0.062	0.084	-0.135	0.102	-0.0941	0.161	-0.515	0.211
2010	-0.032	0.065	-0.126	0.082	-0.0833	0.165	-0.485	0.174
2011	-0.005	0.032	-0.045	0.061	-0.104	0.132	-0.101	0.155
2012	0.0121	0.044	-0.072	0.06	0.005	0.144	-0.252	0.254
2013	0.0125	0.039	-0.065	0.065	0.009	0.152	-0.303	0.281
2014	0.0313	0.03	-0.031	0.0414	0.101	0.177	-0.241	0.211
Average	0.0343	0.0676	-0.0823	0.151	-0.0198	0.1239	-0.2704	0.1813

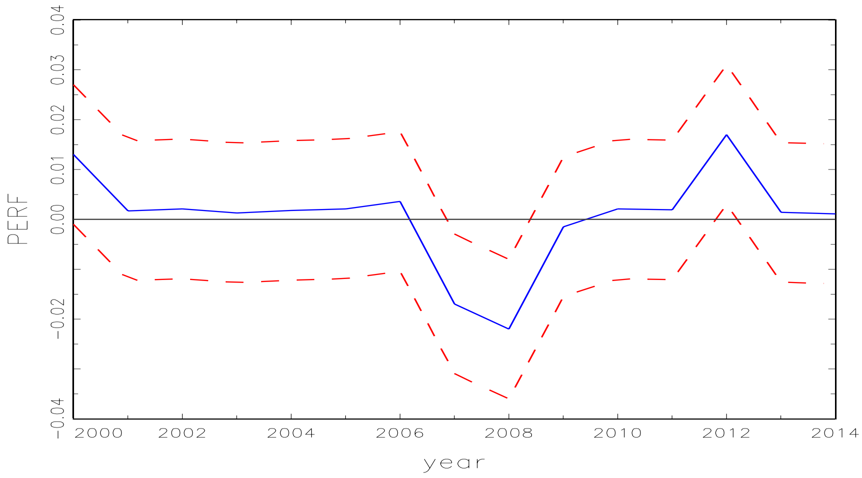
Notes: The table present the yearly PERF measures as in equation (25). Performance is derived from Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration.

Table 9: The yearly average Markowitz performance and persistence indicators over time.

	Performance				Persistence			
Year	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
2000	0.017	0.011	-0.0051	0.022	-0.005	0.012	-0.032	0.014
2001	0.005	0.013	-0.0049	0.021	-0.007	0.011	-0.035	0.021
2002	0.009	0.011	-0.0061	0.024	-0.006	0.01	-0.036	0.023
2003	0.005	0.012	-0.0044	0.026	-0.002	0.009	-0.038	0.018
2004	0.007	0.013	-0.0032	0.032	-0.004	0.015	-0.034	0.028
2005	0.012	0.011	-0.0055	0.029	-0.005	0.019	-0.033	0.036
2006	0.013	0.012	-0.0071	0.029	-0.003	0.021	-0.027	0.039
2007	-0.015	0.013	-0.0265	0.035	-0.002	0.023	-0.038	0.042
2008	-0.022	0.022	-0.0277	0.021	-0.12	0.035	-0.226	0.085
2009	-0.008	0.035	-0.0414	0.009	-0.132	0.042	-0.185	0.072
2010	0.007	0.027	-0.0303	0.013	-0.181	0.037	-0.255	0.072
2011	0.009	0.025	-0.005	0.016	0.004	0.015	-0.044	0.022
2012	0.011	0.022	-0.007	0.015	0.003	0.014	-0.032	0.019
2013	0.012	0.025	-0.006	0.017	0.002	0.012	-0.033	0.015
2014	0.012	0.024	-0.006	0.019	-0.001	0.011	-0.025	0.013
Average	0.0049	0.0184	-0.0124	0.0219	-0.0306	0.0191	-0.0715	0.0346

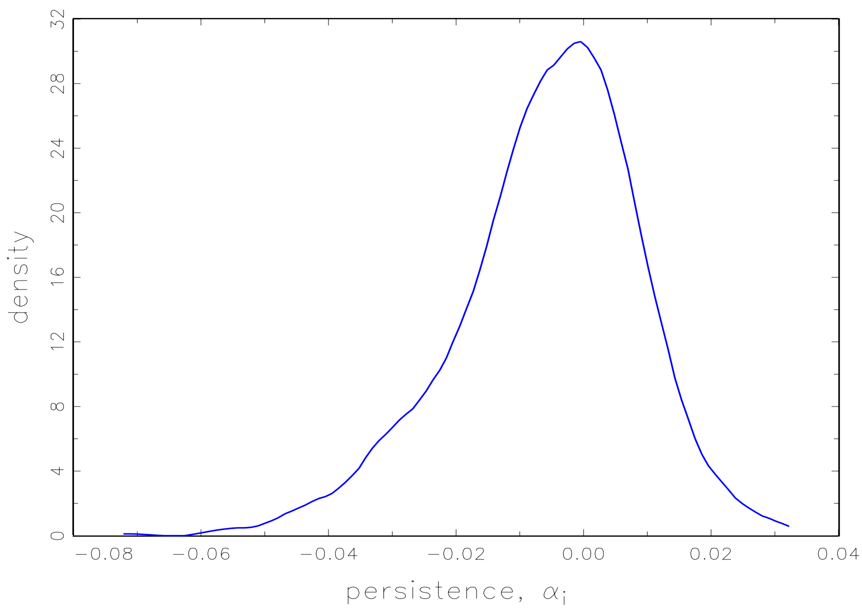
Note: The table presents yearly average Markowitz performance and persistence indicators over time. The posterior distribution is drawn with Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects.

Figure 1: The posterior average overall funds' performance over time.



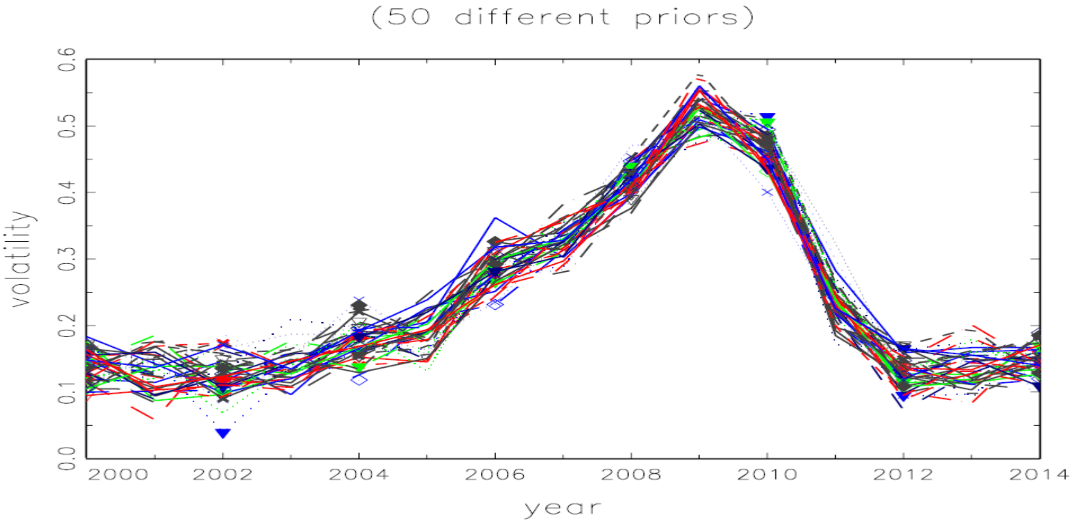
Notes: The figure presents the posterior mean of the overall performance as in equation (25) over time. We have used Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration. Posterior standard deviations appear dashed lines.

Figure 2a: The sample distribution of posterior means of parameters α_i, persistence parameters.



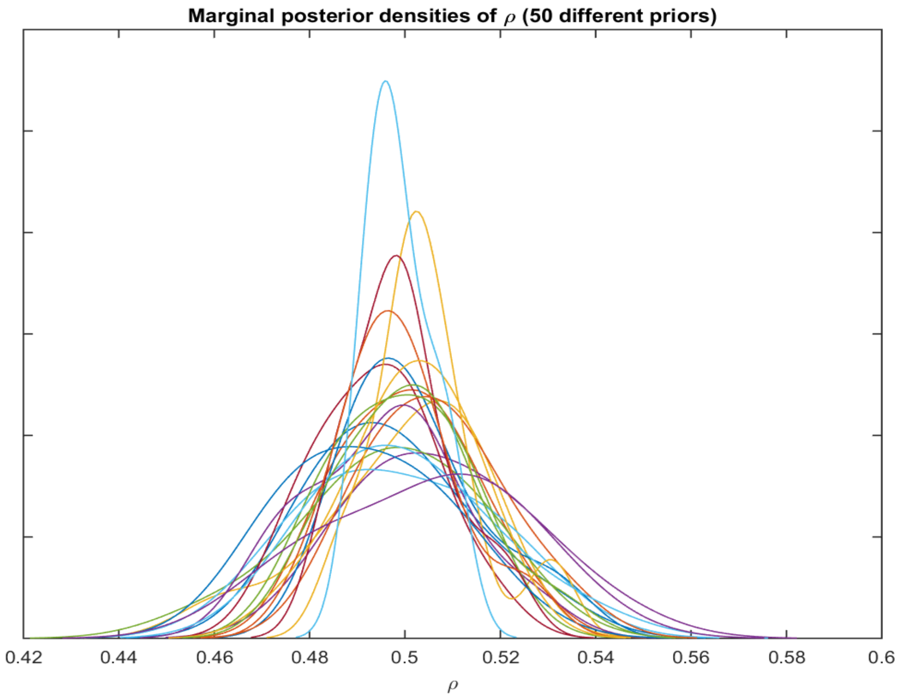
Notes: The figure shows sample distribution of posterior means of parameters α_i, see equation (2). We have used Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration. Posterior standard deviations appear dashed lines.

Figure 2b: The posterior mean funds' volatility ω_i over time average across all funds for 50 different priors.



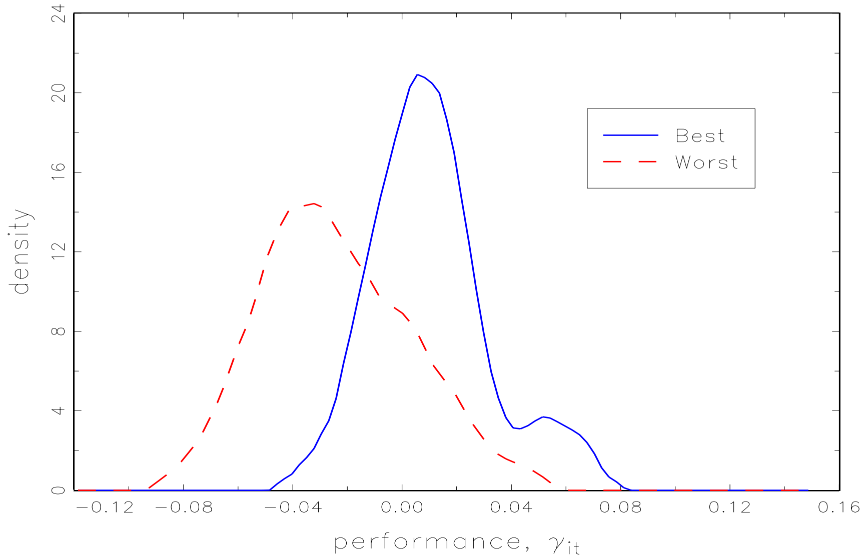
Note: The figure presents the posterior mean funds' volatility ω_i over time average across all funds, see equation (2) for 50 different priors. We have used Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration. Posterior standard deviations appear in dashed lines.

Figure 2c: The posterior densities of ρ corresponding to cross-sectional correlation for 20 different priors.



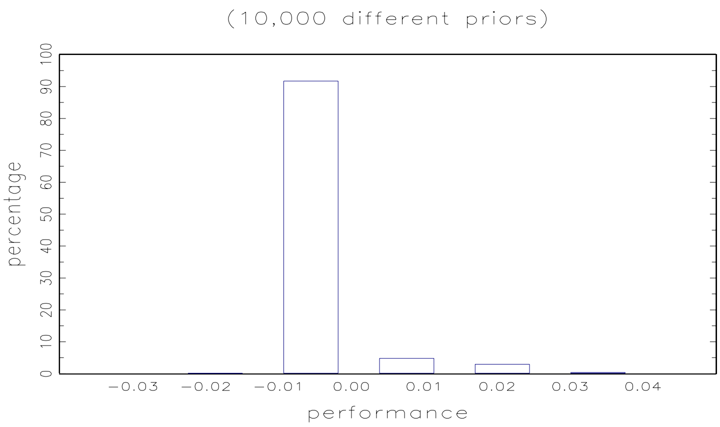
Note: The figure presents the marginal posterior densities of ρ corresponding to cross-sectional correlation for 20 different priors. We have used Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration. Posterior standard deviations appear in dashed lines.

Figure 3: Marginal posterior distribution of g_{it} s



Notes: The figure presents the posterior distribution of g_{it} for the best and the worst performers in the sample. The posterior distribution is drawn with Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects.

Figure 4: funds' performance with all priors.



Note: all 10,000 different priors reported in the last column of Table 1.

APPENDIX I: Sequential Monte Carlo / Particle-Filtering (SMC/PF) techniques.

Chopin (2002) proposed a sequential PF for static models. Given a target posterior $p(\theta | Y) := p(\theta | Y_{1:T})$ a particle system is a sequence $\{\theta_j, w_j\}$ such that $E(h(\theta) | Y) := \int h(\theta) p(\theta | Y) d\theta \cong \lim_{J \to \infty} \frac{\sum_{j=1}^{J} w_j h(\theta_j)}{\sum_{j=1}^{J} w_j}$, almost surely, for any measurable function h, provided the expectation exists. We consider the sequence of posterior distributions $p_t := p(\theta | Y_t)$. The PF algorithm is as follows.

Step 1. Reweight: update the weights $w_j \leftarrow w_j \frac{p_{t+1}(\theta_j)}{p_t(\theta_j)}, j = 1, ..., J$.

Step 2: Resampling: resample $\{\theta_j, w_j J_{j=1}^{H} \rightarrow \{\theta_j^r, 1\}_{j=1}^{J}$.

Step 3. Move: draw $\theta_j^m \sim K_{t+1}(\theta_j^r), j = 1, ..., J$, where K_{t+1} is any transition kernel whose stationary distribution is p_{t+1}.

Step 4. Loop: $t \leftarrow t+1, \{\theta_j, w_j\}_{j=1}^{J} \leftarrow \{\theta_j^m, 1\}_{j=1}^{J}$ and return to Step 1.

Chopin (2002) recommends the independence Metropolis algorithm to select the kernel, which requires a source distribution. A possible choice, if we sampled from p_n ($n < T$), with respect to p_{n+s} is $N(\hat{E}_{n+s}, \hat{V}_{n+s})$ where

$$\hat{E}_{n+s} = \frac{\sum_{j=1}^{J} w_j \theta_j}{\sum_{j=1}^{J} w_j}, \hat{V}_{n+s} = \frac{\sum_{j=1}^{J} w_j \left(\theta_j - E_{n+p}\right)\left(\theta_j - E_{n+p}\right)}{\sum_{j=1}^{J} w_j}.$$

The strategy can be parallelized easily. If K processors are available, we can partition the particle system into K subsets, say $S_k, k = 1, ..., K)$, and implement computations for particles of S_k in processor k. The algorithm can deal with new data at a nearly geometric rate and therefore the frequency of exchanging information between processors (after reweighting) decreases at a rate exponential to n, which is highly efficient.

Resampling according to $\theta_j^m \sim K_t(\theta_j^r,.)$ reduces particle degeneracy (Gilks and Berzuini, 2001) since identical replicates of a single particle are replaced by new ones without altering the stationary distribution. For this application using $J = 2^{12}$ particles gave a mean squared error in posterior means of 10^{-5} over 100 runs.

Chopin (2004) introduces a variation of MSC in which the observation dates at which each cycle terminates (say $t_1,...,t_L$) and the parameters involved in specifying the Metropolis updates (say $\lambda_1,...,\lambda_L$) are specified. Therefore, $0 = t_0 < t_1 < ... < t_L = T$ and we have the following scheme (we rely heavily on Durham and Geweke, 2013).

Step 1. Initialize $l = 0$ and $\theta_{jn}^{(l)} \sim p(\theta)$, $j \in J, n \in N$.

Step 2. For $l = 1,...,L$:

(a) Correction phase:

(i) $w_{jn}(t_{l-1}) = 1, j \in J, n \in N$

(ii) For $s = t_{l-1} + 1,...,t_l$

$$w_{jn}(s) = w_{jn}(s-1) p(y_s \mid y_{1:s-1}, \theta_{jn}^{(l-1)}), j \in J, n \in N.$$

(iii) $w_{jn}^{(l-1)} := w_{jn}(t_l), j \in J, n \in N.$

(b) Selection phase, applied independently to each group $j \in J$: Using multinomial or residual sampling based on $\left\{ w_{jn}^{(l)}, n \in N \right\}$, select

$$\{ \theta_{jn}^{(l,0)}, n \in N \}$$

from $\{ \theta_{jn}^{(l-1)}, n \in N \}$.

(c) Mutation phase, applied independently across $j \in J, n \in N$:

$$\theta_{jn}^{(l)} \sim p(\theta \mid y_{1:t}, \theta_{jn}^{(0)}, \lambda_l) \tag{A1}$$

where the drawings are independent and the pdf above satisfies the invariance condition:

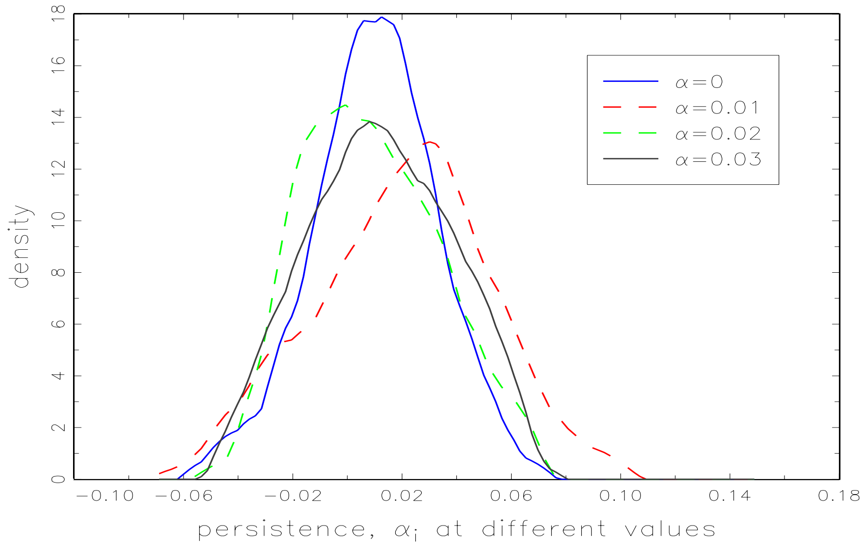
$$\int_{\Theta} p(\theta \mid y_{1:t_l}, \theta^*, \lambda_l) p(\theta^* \mid y_{1:t_l}) d\nu(\theta^*) = p(\theta \mid y_{1:t_l}). \tag{A2}$$

Step 3. $\theta_{jn} := \theta_{jn}^{(l)}, j \in J, n \in N.$

At the end of every cycle, the particles $\theta_{jn}^{(l)}$ have the same distribution $p(\theta \mid y_{1:t_l})$. The amount of dependence within each group depends upon the success of the Mutation phase which avoids degeneracy.

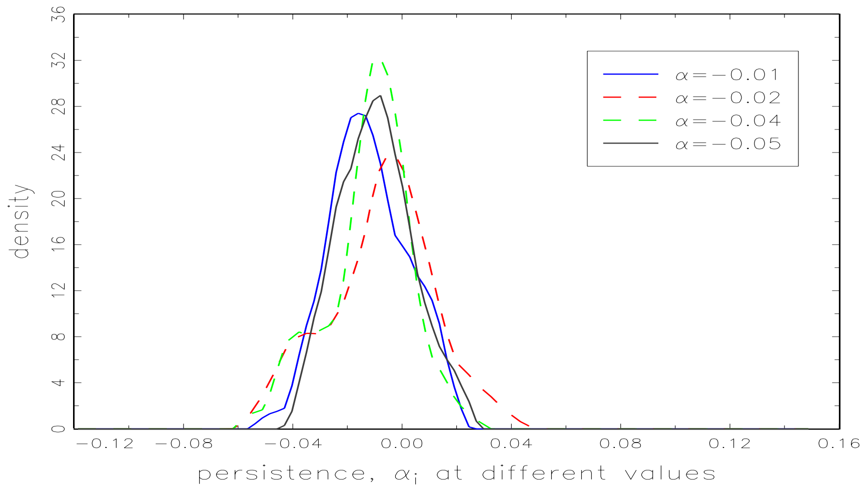
APPENDIX II:

Figure A1a Posterior distributions at different values for $\alpha_i \geq 0$.



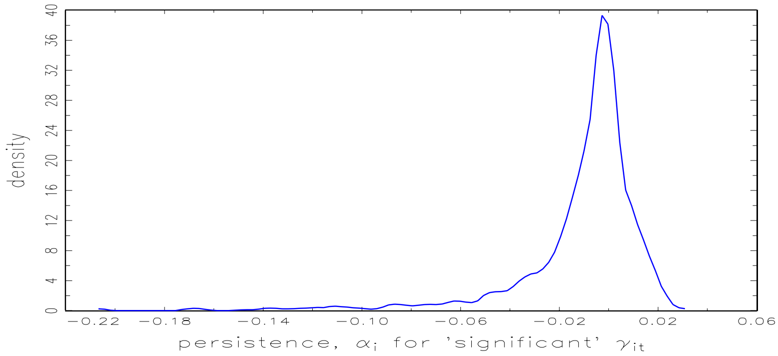
Note: The posterior distribution is obtained with Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects.

Figure A1b Marginal posterior distributions at different values for $\alpha_i \leq 0$.



Note: The posterior distribution is obtained with Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects.

Figure A1c: Sample distribution of posterior mean of α_i conditional on '*significant*' α_i



Notes: The figure presents the sample distribution of posterior mean of α_i conditional on a '*significant*' . The persistence figure is drawn with Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration. Posterior standard deviations appear dashed lines.

Table A1: Yearly 10 average performance indicators over time.

Year	Mean	Standard deviation	Min	Max
2000	0.035	0.041	-0.047	0.120
2001	0.031	0.037	-0.035	0.117
2002	0.033	0.027	-0.048	0.091
2003	0.029	0.032	-0.033	0.097
2004	0.015	0.025	-0.047	0.065
2005	0.019	0.031	-0.032	0.078
2006	0.022	0.035	-0.054	0.094
2007	0.020	0.037	-0.061	0.071
2008	-0.041	0.041	-0.125	0.035
2009	-0.055	0.053	-0.155	0.042
2010	-0.034	0.061	-0.127	0.085
2011	0.012	0.065	-0.110	0.110
2012	0.013	0.075	-0.055	0.118
2013	0.007	0.044	-0.071	0.075
2014	0.011	0.039	-0.075	0.092
Average	0.0078	0.0429	-0.0717	0.0860

Notes: Performance is derived from Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration.

Figure A2a: Posterior distributions of $\overline{\gamma_{it}}$s (performance) : 10 best funds.

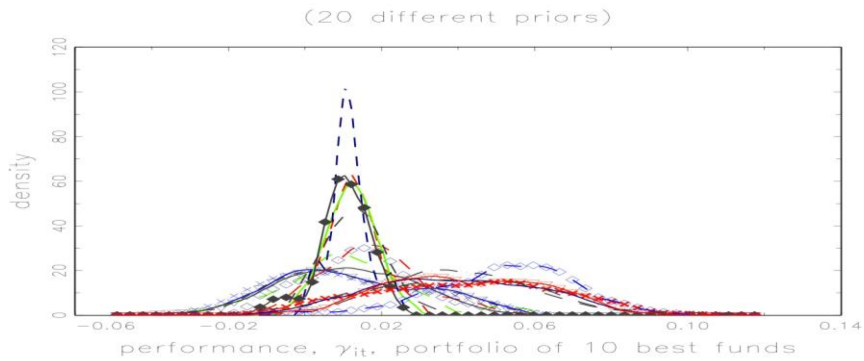


Figure A2b: Posterior distributions of $\overline{\gamma_{it}}$s (performance) : 10 worst funds.

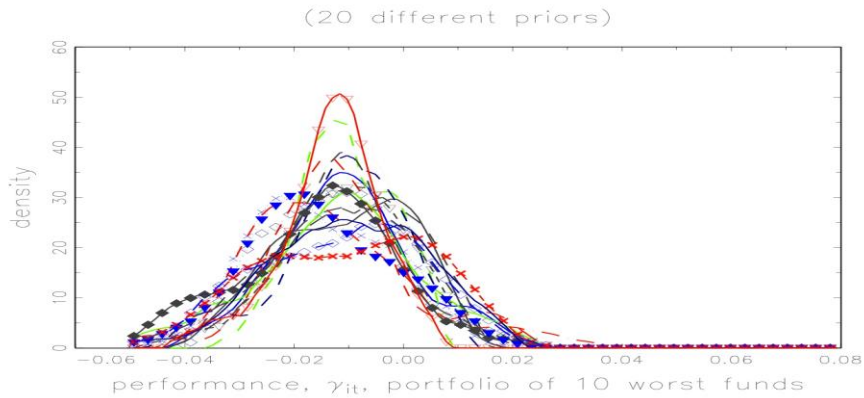


Figure A2c: Posterior distributions of of $\overline{\gamma_{it}}$s (performance): 10 average funds.

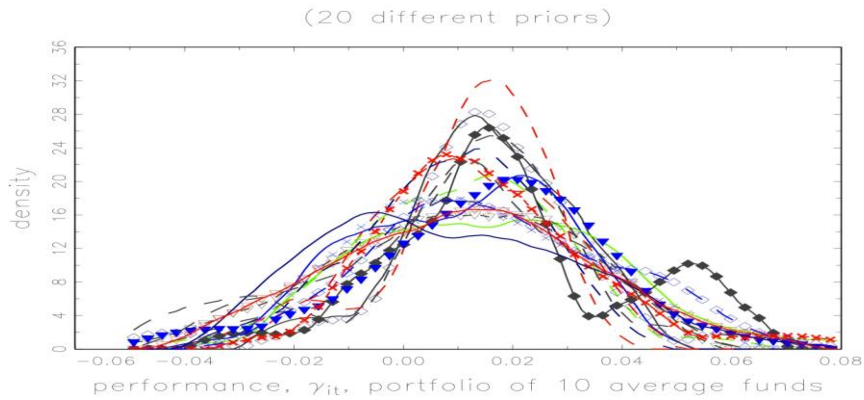
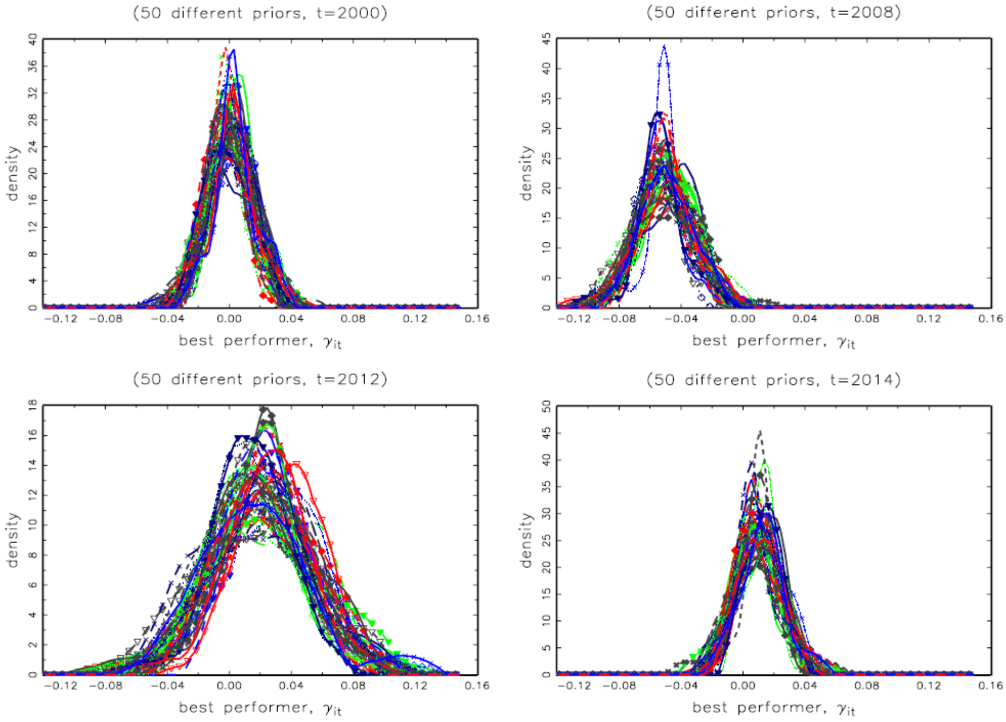
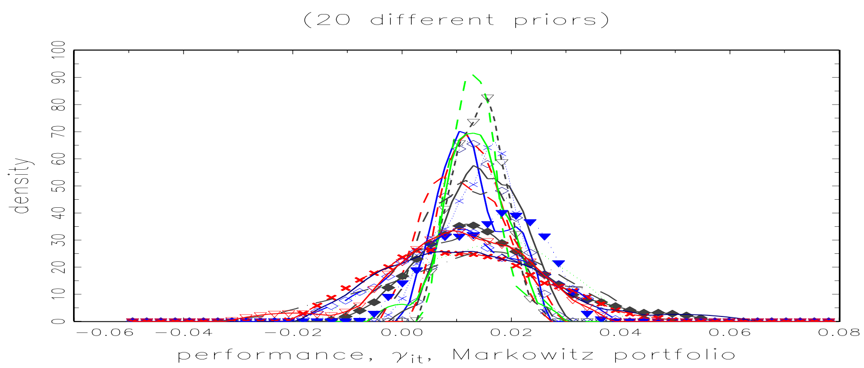


Figure A2d: posterior distributions of γ_{it} for the best-performing fund for years 2000, 2008, 2012 and 2014.



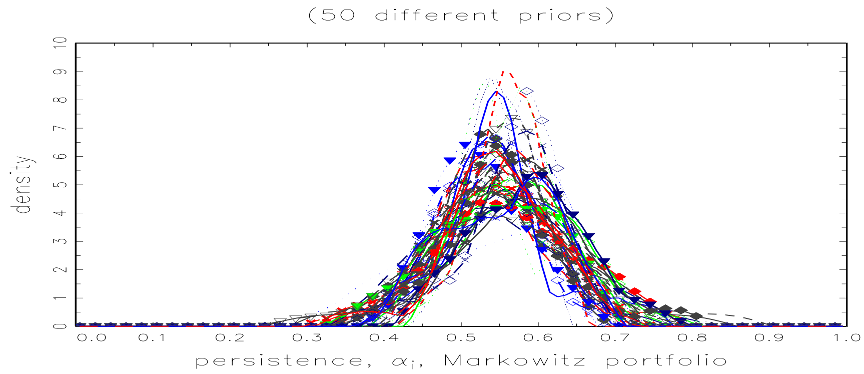
Note: The posterior distributions are obtained with Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects.

Figure A3a: The posterior distribution of optimal portfolio.



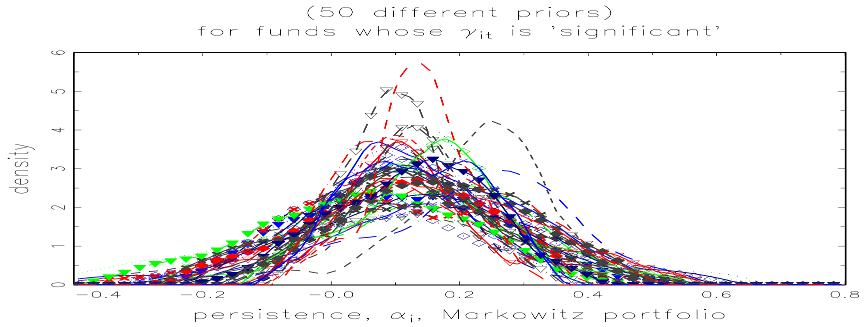
Note: The diagram presents results for 20 different priors from the last column in Table 1.

Figure A3b: Marginal posterior distribution of average persistence ($\overline{\alpha_i}$) of its component returns.



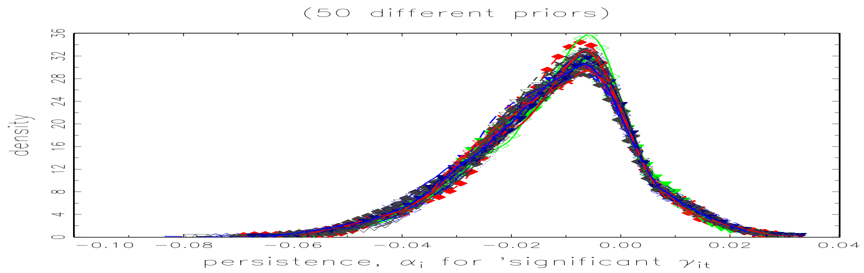
Note: 50 different priors from the last column in Table 1.

Figure A3c: Marginal posterior distribution of persistence when average persistence is '*significant*'.



Note: 50 different priors from the last column in Table 1.

Figure A4: Marginal posterior distributions of α_i for each fund only when its average turns out '*significant*'.



Note: 50 different priors from the last column in Table 1.